6-28-2016 12:00 AM

Characterization of NUMB as a Regulator of Anaplastic Lymphoma Kinase

Ran Wei
The University of Western Ontario

Supervisor
Dr. Shawn Li
The University of Western Ontario

Graduate Program in Biochemistry
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

Cellular events rely on protein-protein interactions that are often mediated by modular domains which recognize particular sequence motifs in binding partners. The NUMB protein is the first described cell fate determinant and multifaceted adaptor that is involved in a wide variety of cellular events. NUMB mainly mediates protein interactions via its modular PTB domain. Here we present a systematic investigation of the NUMB-PTB interactome by employing an integrative strategy combining both protein and peptide arrays. We profiled NUMB-PTB binding specificity and interacting proteins genome-wide. The receptor tyrosine kinases (RTKs) are found highly enriched in the interactome, raising the possibility that NUMB may act as a universal binding partner of RTKs. To further validate this hypothesis, we focused on the interaction between NUMB and Anaplastic Lymphoma Kinase (ALK), which promotes oncogenesis in a number of cancer types. Consistent with the prediction based on our proteomic study, NUMB-PTB directly binds to two motifs in ALK *in vitro* and *in vivo*. Intriguingly, functional analysis reveals that NUMB-ALK interaction regulates ALK activity antagonistically in an isoform dependent manner, by directing ALK to distinct post-endocytic trafficking destinations. Our study provides mechanistic insight into ALK regulation, explains the controversial behaviors of NUMB in tumorigenesis at the molecular level, and further reveals a biomarker of potential clinical value.

Keywords

NUMB, PTB, interactome, binding motifs, isoform, receptor tyrosine kinase, ALK, endocytosis, post-endocytic trafficking, recycling, degradation, MAPK, proliferation, cancer

Co-Authorship Statement

All chapters of this thesis were written by Ran Wei and edited by Dr. Shawn Li and Ms. Courtney Voss. Most experimental results and analysis presented in this thesis were completed by Ran Wei.

The data presented in Chapter 2 is in preparation for publication. The chip-based NUMB interactome screening was assisted by Dr. David Schibli, Dr. Chengang Wu. The bioinformatics data analysis was assisted by Dr. Lei Li and Dr. Huadong Liu. The structural analysis was assisted by Dr. Tomo Kaneko.

The data presented in Chapter 3 is in preparation for publication. All confocal microscopy, image processing and quantification were assisted by Xuguang Liu.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Shawn Li for providing me with this great opportunity to study at Western and work on fabulous research projects. In the past years, Dr. Li has provided me continuous support and helped me a lot in and outside the laboratory. Dr. Li's expertise assisted me in establishing a firm foundation in research. I could not wish for a better supervisor for my graduate studies.

I would like to thank my committee Dr. David Litchfield, Dr. Eric Ball and Dr. Lina Dagnino for their invaluable suggestions for developing my project. Many thanks go to Dr. Marc Vigny for his insightful comments and his generous permission to share a number of critical experimental materials. I also would like to thank Ms. Barb Green for her continuous administrative service.

Thank you all to Li lab members. I would like to give my special thanks and appreciation to Ms. Courtney Voss, Dr. Huadong Liu, Dr. Tomonori Kaneko, Dr. Shelley Sandiford and Dr. Gurpreet Dhami. My thanks also go to Dr. Lei Li, Dr. Mei Huang, Dr. Rakesh Joshi, Dr. Kyle Biggar, Dr. Lyugao Qin, Dr. Xing Li, Xuguang Liu, Qi Fang, Eric Liu and Wayne Wu. Because of you, Li lab is such an adorable and enjoyable place to work and study.

Finally, I would like to thank my parents Guoqiang Wei and Xiaoqin Zhou for their understanding and encouragement.

Table of Contents

Abstract i

Keywords ii

Co-Authorship Statement iii

Acknowledgments iv

Table of Contents v

List of Figures and Tables ix

List of Abbreviations xi

Chapter 1 1

1 General introduction 1

1.1 The multifaceted adaptor protein NUMB 1

1.1.1 NUMB structure and isoforms 2

1.1.2 NUMB in cancer as a tumor suppressor 5

1.1.3 The controversial functions of NUMB in cancer 11

1.2 Receptor tyrosine kinases in cancer 13

1.2.1 Tyrosine phosphorylation 13

1.2.2 Tyrosine phosphorylation mediated cell signaling 15

1.2.3 Endocytosis regulates RTK signaling 20

1.2.4 ALK-positive cancers and ALK target therapeutics 23

1.3 Scope of thesis project 26

1.4 Reference 28

Chapter 2 39

2 Investigation of NUMB-PTB interactome 39

2.1 Abstract 39

2.2 Introduction 40

2.3 Material and methods 44

2.3.1 Cell culture 44
2.3.2 GST-NUMB-PTB expression, purification and biotinylation 44
2.3.3 ProtoArray hybridization 45
2.3.4 Transient transfection 45
2.3.5 Immunoprecipitation, pull down and western blotting 45
2.3.6 Free peptide and peptide array synthesis 46
2.3.7 Far western assay for peptide membrane blotting 46
2.3.8 Determination of dissociation constant 47
2.3.9 Consensus binding motif and SMALI matrix 47
2.4 Results 49
2.4.1 Screening NUMB PTB domain binding proteins using protein array 49
2.4.2 Characterizing NUMB-PTB domain binding sites in the interactome 51
2.4.3 Defining NUMB-PTB domain binding specificity 56
2.4.4 Structural explanation of NUMB-PTB binding specificity 57
2.4.5 Verification of NUMB RTK interactors in vivo 60
2.4.6 Predicting a genome-wide NUMB interactome by SMALI 61
2.5 Discussion 63
2.5.1 An integrative array strategy for systematic investigation of domain-peptide interaction mediated PPI network. 63
2.5.2 The genome-wide prediction of NUMB-PTB binding partners 63
2.5.3 NUMB may act as a universal regulator to RTKs 65
2.6 Reference 66
Chapter 3 71
3 NUMB regulates ALK endocytosis and activity in an isoform-dependent manner . 71
3.1 Abstract 71
3.2 Introduction 72
3.3 Material and methods 75
3.3.1 Cell culture 75
3.3.2 Transient transfection 75

3.3.3 siRNA interference 75
3.3.4 Immunoprecipitation and western blotting 75
3.3.5 Recombinant protein expression and GST pull-down 76
3.3.6 Gene truncation and site-directed mutagenesis 76
3.3.7 Immunostaining and microscopic quantification 77
3.3.8 Peptide synthesis 78
3.3.9 Determination of dissociation constant 78
3.3.10 Cell surface biotinylation assay and recycling assay 79
3.3.11 Soft agar assay 79
3.3.12 Cell proliferation assay 80
3.4 Results 81
3.4.1 Identification of ALK as a novel NUMB-binding protein 81
3.4.2 NUMB regulates ALK endocytosis 84
3.4.3 NUMB regulates ALK post-endocytic trafficking 85
3.4.4 NUMB regulates ALK activity via mediating ALK endocytosis 91
3.5 Discussion 95
3.6 References 98
Chapter 4 102
4 ALK ectodomain shedding 102
4.1 Abstract 102
4.2 Introduction 103
4.3 Material and methods 105
4.3.1 Cell Culture and treatment 105
4.3.2 PCR mediated ALK truncations 105
4.3.3 Immunoprecipitation and western blotting 106
4.3.4 Protein digestion and mass spectrometry analysis 107
4.4 Results 108
4.4.1 Identification of ALK-140-kDa C-terminal fragment 108
4.4.2 Screening for ALK cleavage site by MS 109

4.4.3 MMP and ADAM mediate ALK ectodomain shedding 110
4.4.4 140-kDa ALK fragment is highly phosphorylated 112
4.5 Discussion 113
4.6 References 114
Chapter 5 117
5 Summary and perspectives 117
5.1 Summary of study 117
5.2 NUMB isoforms, endosomal sorting and cancer diagnosis 121
5.3 References 123
Supplemental Data 128
Curriculum Vitae 172

List of Figures and Tables

Figure.1.1 Structure of mammalian NUMB isoforms 3

Figure.1.2 NUMB-regulated NOTCH signal transduction 8

Figure.1.3 Control of p53 activity by NUMB and MDM2 10

Figure.1.4 Three-part toolkit of tyrosine phosphorylation 14

Figure.1.5 Functions of select p-Tyr sites in the EGFR receptor 18

Figure.1.6 Oncogenic forms of ALK 25

Figure.2.1 A strategy to identify NUMB-PTB binding specificity and interactome 42

Figure.2.2 NUMB PTB domain interactome 50

Figure.2.3 Verifying NUMB-PTB domain binding sites in the interactome 52

Figure.2.4 Binding curves for a selection of NUMB-PTB interactors 54

Figure.2.5 Amino acid preferences of verified NUMB-PTB binding motifs 55

Figure.2.6 Identifying NUMB-PTB binding specificity 57

Figure.2.7 Structural analysis of NUMB-PTB binding specificity 59

Figure.2.8 NUMB binds to ALK, ErbB2, and FGR receptors *in vivo* 60

Figure.2.9 Prediction of NUMB-PTB binding motifs genome-wide 62

Figure.2.10 NUMB binds to motifs in PTP-9 and PTP-13 62

Figure.3.1 Identification of the interaction between NUMB and ALK 82

Figure.3.2 NUMB binds two motifs in ALK 83

Figure.3.3 NUMB promotes ALK internalization 85

Figure.3.4 NUMB isoforms play distinct roles in ALK post-endocytic trafficking..... 88

Figure.3.5 NUMB isoforms differentiate the co-localization between ALK and Rabs.... 90

Figure.3.6 p66-NUMB inhibits ALK-mediated MAPK signaling 92

Figure.3.7 p66-NUMB inhibits ALK-dependent cancer cell growth 94

Figure.3.8 Overview of NUMB-regulated ALK endocytosis..... 96

Figure.4.1 Identification of ALK-140-kDa fragment 108

Figure.4.2 MS analysis of ALK ectodomain 109

Figure.4.3 MMPs mediate ALK ectodomain shedding 111

Figure.4.4 140-kDa ALK fragment is highly phosphorylated..... 112

S1 ProtoArray grid design 128

S2 ProtoArray quantification of NUMB-PTB domain (Z-Score>3.0) 129

S3 NUMB interactome 135

S4 NUMB-PTB peptide array quantification 138

S5 OPAL Scoring Matrix of NUMB-PTB..... 141

S6 Prediction of NUMB-PTB binding partners..... 142

S7 MS analysis of ALK ectodomain (chymotrypsin digestion) 169

S8 MS analysis of ALK ectodomain (Glu-C digestion) 170

S9 MS analysis of ALK ectodomain (Trypsin digestion)..... 171

List of Abbreviations

Abbreviation	Full Name
ADAM	a disintegrin and metalloproteinase
ALCL	anaplastic large cell lymphomas
ALK	anaplastic lymphoma kinase
AP-MS	affinity-purification mass spectrometry
bHLH	basic helix-loop-helix DNA binding domain
Co-IP	co-immunoprecipitation
ConA	concanamycin A
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulfoxide
dNUMB	Drosophila NUMB
DPF	Asp-Pro-Phe motif
DSL	Delta/Serrate/LAG-2 proteins
DTT	dithiothreitol
ECL	enhanced chemiluminescence
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
EGTA	ethylene glycol tetraacetic acid
EMT	epithelia-mesenchymal transition
ErbB2	Erythroblastic leukemia viral oncogene
Erk	extracellular signal-regulated kinase
F1174L	ALK constitutive activation mutate F1174L
F162V	NUMB PTB binding deficient mutate F162V
FBS	fetal bovine serum
FGR	Gardner-Rasheed feline sarcoma viral oncogene homolog
FLAG	FLAG octapeptide

Fmoc	9-fluorenylmethyloxycarbonyl
FP	fluorescence polarization
GFP	green fluorescent protein
GSH	glutathione
GST	glutathione s-transferase
HA	hemagglutinin
HER2	Human epidermal growth factor 2
HES	hairy/enhancer-of-slit
hNUMB	human NUMB
hNUMBL	human NUMB-like
HRT/HEY	hair-related transcription factor
HRT/HEY	hairy-related transcription factor
IPTG	isopropyl β-D-a-thiogalactopyranoside
JACoP	just another colocalization plugin of ImageJ software
Kd	dissociation constant
LB	lysogeny broth medium
LNX	ligand of NUMB-protein X
LUMIER	luminescence-based mammalian interactome mapping
MAML	MASTERMIND-like transcription activators
MAPK	mitogen-activated protein kinase
MDM2	mouse double minute 2 homolog
MEM	minimum essential media
MMP	matrix metalloproteinase
N1477A	ALK PTB binding deficient mutate N1477A
N1583A	ALK PTB binding deficient mutate N1583A
NGF	nerve growth factor
NICD	NOTCH intercellular domain
NOTCH	neurogenic locus notch homologue protein
NPF	Asn-Pro-Phe motif

NSCLC	non small cell lung cancer
OPAL	oriented peptide array library
p21	CIP1/WAF1 cyclin-dependent kinase inhibitor
p65	NUMB isoform 65 kDa
p66	NUMB isoform 66 kDa
p71	NUMB isoform 71 kDa
p72	NUMB isoform 72 kDa
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction of DNA amplification
pen/strep	penicilin/streptomycin antibiotics
PI3K	phosphatidylinositol-4,5-bisphosphate 3-kinase
PKC	protein kinase c
PMA	Phorbol-12-Myristate-13-Acetate
PPI	protein-protein interaction
PRR	proline rich region
PTB	phosphotyrosine binding domain
PTBi/PTBS	NUMB phosphotyrosine binding domain short form
PTBo/PTBL	NUMB phosphotyrosine binding domain full length
PTK	protein tyrosine kinase
PTM	post-translational modification
PTP	protein tyrosine phosphatase
pTyr	phosphotyrosine
PVDF	polyvinylidene difluoride
R1275Q	ALK constitutive activation mutate R1275Q
Rab4DN	dominate negative mutate of Rab4 protein
Rab7DN	dominate negative mutate of Rab7 protein
RBPJ	recombining binding protein suppressor of hairless

ROI	regions of interest in imaging processing
ROS	proto-oncogene tyrosine-protein kinase
RPMI	Roswell Park Memorial Institute
RTK	receptor tyrosine kinase
SDS	sodium dodecyl sulfate
Set8	N-lysine methyltransferase SETD8
SH2	Src homology 2
siRNA	small interference RNA
SMALI	matrix-assisted ligand identification
T-ALL	T-cell acute lymphoblastic leukemia
TCRβ	T-cell receptor beta
TFA	trifluoroacetic acid
TIPS	tri-isopropylsilane
TK	tyrosine kinase
TKI	tyrosine kinase inhibitor
Tris	trisaminomethane
TrkB	tropomyosin receptor kinase B
WT	wild type

Chapter 1

1 General introduction

1.1 The multifaceted adaptor protein NUMB

NUMB is the first characterized cell fate determinant, originally described in *Drosophila* in 1989, where it plays a role in controlling the fate of progeny derived from sensory organ precursors (Uemura et al., 1989). Later in 1994, NUMB was reported to regulate cell fate through asymmetrical distribution in daughter cells during sensory organ precursor division (Rhyu et al., 1994). NUMB is universally distributed at the membrane periphery during interphase, but preferentially accumulates at one spindle pole during mitotic division and thus is inherited by only one daughter cell which then develops a distinct cell fate. It subsequently became clear that NUMB directly binds to and counteracts the membrane receptor NOTCH (neurogenic locus notch homologue protein) in this cell lineage, since NOTCH gain-of-function mutations matched the phenotype of NUMB loss-of-function mutations (Guo et al., 1996). The report of this novel yet significant NUMB function has henceforth captivated the attention of researchers.

The NUMB gene is evolutionarily conserved from fly to human. *Drosophila* NUMB (dNUMB) and mammalian NUMB (mNUMB) proteins are also functionally conserved, since ectopic expression of mNUMB complemented the phenotype of dNUMB knockout in *Drosophila* (Salcini et al., 1997; Zhong et al., 1996; Zhong et al., 1997). In general, the NUMB protein exhibits a complex pattern of functions such as the control of asymmetric cell division and cell fate, endocytosis, cell adhesion, cell migration, ubiquitination and several signaling pathways (Gulino et al., 2010). The complexity of NUMB is likely due to multiple isoforms generated by alternative splicing, and their diverse expression and subcellular localization patterns. In humans, studies on NUMB mainly focus on its role in tumorigenesis and have shown the involvement of NUMB in a variety of cancer-related events and signaling pathways (Gulino et al., 2010; Pece et al., 2011). In particular, NUMB has been defined as a tumor suppressor by counteracting the function of NOTCH, one of

the major oncogenic receptors in many different tumor types (Espinoza and Miele, 2013; Leong and Gao, 2008), or by stabilizing p53 to attenuate the immortality of cancer cells (Colaluca et al., 2008).

1.1.1 NUMB structure and isoforms

In different species, the NUMB proteins share a similar structure as presented in Fig.1.1. Mammalian NUMB contains an N-terminal phosphotyrosine binding (PTB) domain, a proline rich region (PRR), as well as two Asp-Pro-Phe (DPFs) motifs and one Asn-Pro-Phe (NPF) motif at the C-terminus (Bork and Margolis, 1995; Li et al., 1997; Li et al., 1998; Zwahlen et al., 2000). The PTB domain and PRR are the major protein interaction sites in NUMB, and contribute to most NUMB functions that have been characterized to date. The PTB domain is crucial for the membrane association of NUMB, because it directly interacts with a number of membrane localized/associated proteins (Dho et al., 1999; Dho et al., 2006; Gulino et al., 2010), such as the receptor tyrosine kinases ErbB2 and ALK discussed in Chapter 2. These interactions assimilate NUMB into various signaling transduction pathways initiated from the cell membrane. The PRR, as well as the DPF and NPF motifs, directly interact with various endocytic adaptor proteins such as α-adaptin and Eps15 (Salcini et al., 1997; Santolini et al., 2000). This interaction pattern implicates the potential role of NUMB in endocytosis regulation. Indeed, NUMB has been shown to regulate the endocytosis of various membrane localized receptors thus indirectly altering their downstream signaling transduction (Gulino et al., 2010).

In both mice and humans, NUMB undergoes alternative splicing to produce at least 9 isoforms based on the Uniprot database. Alternative splicing occurs in exon 3 in the PTB domain or exon 9 in the PRR to generate the four major isoforms: p72 (inclusion of both exons), p71 (inclusion of exon 9), p66 (inclusion of exon 3) and p65 (exclusion of both exons) (Dho et al., 1999; Verdi et al., 1999). Five shorter isoforms have also been reported, but their characteristics and functions are not well studied.

Although the diverse biological effects of the four major isoforms have been reported in a number of studies, the functional mechanisms of these isoforms have not been well

elucidated. It is speculated that the diversity is mainly caused by differences in protein binding specificities among the long (L) and short (S) variants of PTB domain and PRR.

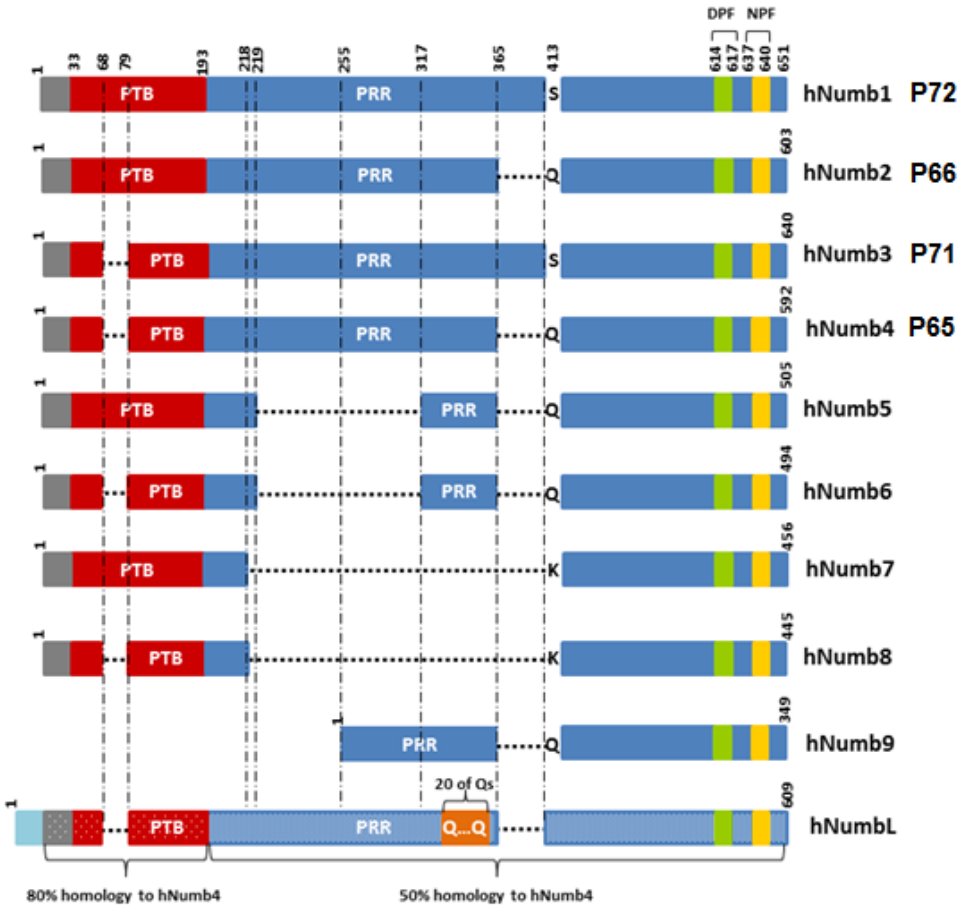


Figure.1.1 Structure of mammalian NUMB isoforms

Overview of NUMB protein isoforms in human. The human genome encodes two NUMB homologs, NUMB and NUMB-like (NUMBL). NUMB undergoes alternative splicing to generate at least 9 isoforms, including 4 major isoforms that have been characterized. The four isoforms (hNUMB1-p72 isoform; hNUMB2-p66 isoform; hNUMB3-p71 isoform; hNUMB4-p65 isoform) are differentiated by alternative splicing in the PTB domain (exon 3) and PRR (exon9). NUMBL only expresses one protein, which is similar to the hNUMB4-p65 isoform and contains the short forms of both PTB domain and PRR.

A small 11 amino acid insert in the PTB domain distinguishes p72-NUMB/p66-NUMB (PTB^L, or PTBi) from p71-NUMB/p65-NUMB (PTB^S, or PTBo) (sequence: ERKFF KGFFG K). This variation in the PTB domain is believed to determine the subcellular localizations of NUMB isoforms. The PTB^L isoforms, p72 and p66, tend to localize at the

cell membrane periphery, PTB^S isoforms, p71 and p65, are predominantly cytosolic (Dho et al., 1999a). A binding partner of NUMB, the Ligand of NUMB-protein X (LNX), which is an E3 ubiquitin ligase, is reported to ubiquitinate PTB^L isoforms but not PTB^S isoforms (Dho et al., 1998; Nie et al., 2004; Nie et al., 2002). Other than these examples of functional diversity, few differences have been described between the PTB^L and PTB^S isoforms. As will be presented in Chapter 3, the PTB^L and PTB^S isoforms exhibit no differences in binding to NxxF and NxxY motifs or ALK receptor tyrosine kinase both *in vitro* and *in vivo*.

A large 48 amino acid insert in the PRR distinguishes p72-NUMB/p71-NUMB (PRR^L, or PRRi) from p66-NUMB/p65-NUMB (PRR^S, or PRRo) (sequence: ANGTD SAFHV LAKPA HTALA PVAMP VRETN PWAHA PDAAN KEIAA TCS). In preliminary studies, the PRR^L and PRR^S variants exhibited distinct functional patterns in different models of development. *In vitro*, the PRR^S isoforms promote cell differentiation and are mainly expressed in the early stages of neural crest stem cell development, while the PRR^L isoforms promote cell proliferation and are expressed throughout the entire developmental process (Verdi et al., 1999). *In vivo*, an expression pattern switch was observed during mouse neuronal development. Expression of PRR^L or PRR^S isoforms was prominent before and after the cortical progenitor differentiation (Bani-Yaghoub et al., 2007). However, only the PRR^L isoforms were detected consistently in adult mouse brain tissue (Dho et al., 1999). A similar effect was also observed in mouse pancreatic and rat retinal development, in which overexpression of PRR^L isoforms coincides with an expanding progenitor population (Dooley et al., 2003; Yoshida et al., 2003). Although the NUMB gene only expresses one protein without alternative splicing in *Drosophila,* ectopic expression of human PRR^L and PRR^S isoforms results in proliferation and differentiation of neural stem cells respectively (Toriya et al., 2006).

Nevertheless, the mechanism of this isoform-dependent functional diversity has not been well elucidated at the molecular level. It is speculated that this diversity correlates with distinct interactomes for PRR^L and PRR^S isoforms in specific genetic backgrounds. In Chapter 3, a model will be presented that will partially explain a functional difference

between NUMB PRR^L and PRR^S isoforms. The isoforms are found to direct a receptor tyrosine kinase to different destinations during endosomal sorting.

1.1.2 NUMB in cancer as a tumor suppressor

Despite the important function of NUMB in the neural system and brain development, current research in human NUMB is centered on studying its role in cancer biology. It is not surprising that this cell fate determinant is involved in tumorigenesis. Previously, a conceptual correlation between cancer and asymmetric cell division was established (Morrison and Kimble, 2006; Reya et al., 2001). Stem cells are defined by their ability to generate more stem cells, known as self-renewal, as well as differentiating cells, usually via asymmetric cell division. A number of cell fate determinants, including NUMB, are involved in the regulation of asymmetric cell division. Impairment of this process can result in an imbalance between self-renewal and differentiation, resulting in disordered cell division and proliferation, which are key characteristics of cancer cells. Along with the correlation between cancer and asymmetric division, NUMB has been defined as a tumor suppressor since it attenuates the activity of the NOTCH receptor, one of the most frequently deregulated receptors in solid tumors (Flores et al., 2014), and assists in the stabilization of p53 to increase apoptotic susceptibility of immortal cancer cells (Colaluca et al., 2008).

1.1.2.1 NUMB regulates the NOTCH signaling pathway

The NOTCH (neurogenic locus notch homologue protein) pathway is an evolutionarily conserved signaling pathway present in most multicellular organisms, which has received increased recognition as a major deregulated signaling pathway in many solid tumors (Espinoza and Miele, 2013; Leong and Gao, 2008). In tumorigenesis, the pivotal function of NOTCH is similar to that of the NUMB protein, which controls cell fate determination and cellular proliferation (Artavanis-Tsakonas et al., 1999; Chiba, 2006). At the tissue level, NOTCH is also involved in angiogenesis and epithelial-mesenchymal transition (EMT) (Leong and Karsan, 2006).

The human NOTCH family consists of four homologs (NOTCH1-4). NOTCH proteins are expressed as full-length unprocessed proteins composed of a large extra cellular domain, a single transmembrane domain, and an intracellular domain. During transportation from the secretory system to cell membrane, NOTCH proteins are cleaved to generate two subunits: one consisting of the majority of the extracellular domain, and the other consisting the remaining extracellular domain, the transmembrane domain and the intracellular domain (Blaumueller et al., 1997; Capobianco et al., 1997). The two subunits finally associate via a calcium-dependent, noncovalent interaction to form a mature heterodimeric transmembrane receptor.

NOTCH proteins are activated by a transmembrane ligand family of DSL proteins (Delta and Serrate/Jagged in *Drosophila* and mammals, LAG-2 in *Caenorhabditis elegans*) with five members in mammals: Jagged 1/2 and Delta-like 1/3/4, whose extracellular regions serve as NOTCH ligands (Nam et al., 2003). Upon ligand binding, the NOTCH extracellular domain undergoes a conformational change, resulting in the exposure of a cleavage site. ADAM (a disintegrin and metalloproteinase) family proteins Kuzbanian (ADAM10) and TACE (ADAM17) shed the extracellular domain at this site, followed by another intracellular cleavage mediated by γ-secretase to release the entire intracellular fragment of NOTCH (NICD, NOTCH intracellular domain) (Bray, 2006; Brou et al., 2000; Wilkin and Baron, 2005). The NICD functions as a transcription co-activator after translocation into the nucleus, where it interacts with CSL nuclear transcription factor (CBF1/RBPJ, recombining binding protein suppressor of hairless), or the family of MASTERMIND transcription activators (MAML1/2/3) to promote the expression of genes required for the execution of NOTCH pleiotropic functions (Brou, 2009; Davis and Turner, 2001; Iso et al., 2003). In mammals, the primary targets of NOTCH are genes belonging to the basic helix-loop-helix (bHLH) DNA binding proteins (Iso et al., 2003), containing at least two families: the hairy/enhancer-of-slit (HES) family and the hairy-related transcription factor (HRT, also known as HEY) family. Because both the NOTCH receptors and the DSL ligands are transmembrane proteins, activation of the NOTCH pathway is intercellular and directional from signal sending (DSL expressing) cells to signal receiving (NOTCH expressing) cells.

In healthy tissue, the best-characterized function of NOTCH is the determination of cell fate (Artavanis-Tsakonas et al., 1999; Chiba, 2006; Lai, 2004). Similarly to NUMB, aberrant NOTCH signaling has been implicated in human cancer (Roy et al., 2007). NOTCH was first identified as an oncogene in T-cell acute lymphoblastic leukemia (T-ALL) in the form of chromosomal translocation fusion protein containing the N-terminal region of the T-cell receptor beta (TCRβ) and C-terminal of NOTCH (Ellisen et al., 1991). Identical to many other translocation fusion proteins of membrane receptors, this leads to expression of a constitutively active form of NOTCH, which continuously stimulates downstream signaling (Greenwald, 1994). It was later discovered that deregulation of NOTCH signaling is potentially oncogenic in a multitude of solid tumors and several other types of leukemia as well (Espinoza and Miele, 2013; Pancewicz and Nicot, 2011). To date, this list has rapidly expanded and contains almost all major solid cancer types, including cervical, head and neck, endometrial, renal, lung, pancreatic, ovarian, prostate, esophageal, oral, hepatocellular and gastric carcinomas, osteosarcoma, mesothelioma, melanoma, gliomas, medulloblastomas and rhabdomyosarcoma (Espinoza and Miele, 2013).

The mechanisms of NOTCH signaling over-activation vary, but deregulation by NUMB loss-of-function plays a pivotal role in this process. In fact, altered NOTCH signaling (by intact NOTCH gene) is diagnosed in approximately one third of non-small-cell lung cancer (NSCLC) cases, which are correlated with wild type p53 and NUMB deletion. In contrast, gain-of-function NOTCH mutations are only identified in a much smaller fraction (around 10%) of NSCLC cases (Westhoff et al., 2009). The NUMB-NOTCH interplay is well-recognized as a tumor suppressor, however the precise mechanism continues to be reappraised based on new evidence. To date, several models have been proposed based on isolated reports. NUMB promotes NOTCH endocytosis and lysosomal degradation to attenuate NOTCH protein level in cells (Barth and Köhler, 2014; McGill et al., 2009; Yamamoto et al., 2010). In addition, NUMB increases NOTCH ubiquitination and degradation (McGill and McGlade, 2003), but it remains uncertain whether this ubiquitination takes place during the endocytosis of NOTCH, a situation that is common in the regulation of many membrane receptors. Other evidence also supports that NUMB binding prevents nuclear transportation of the NICD domain following full activation and cleavage of NOTCH (Berdnik et al., 2002; Frise et al., 1996; Gho et al., 1996). The models

of NOTCH activation, signal transduction, NUMB regulation are illustrated in Fig.1.2, and recently reviewed by Flores *et al* (Flores et al., 2014).

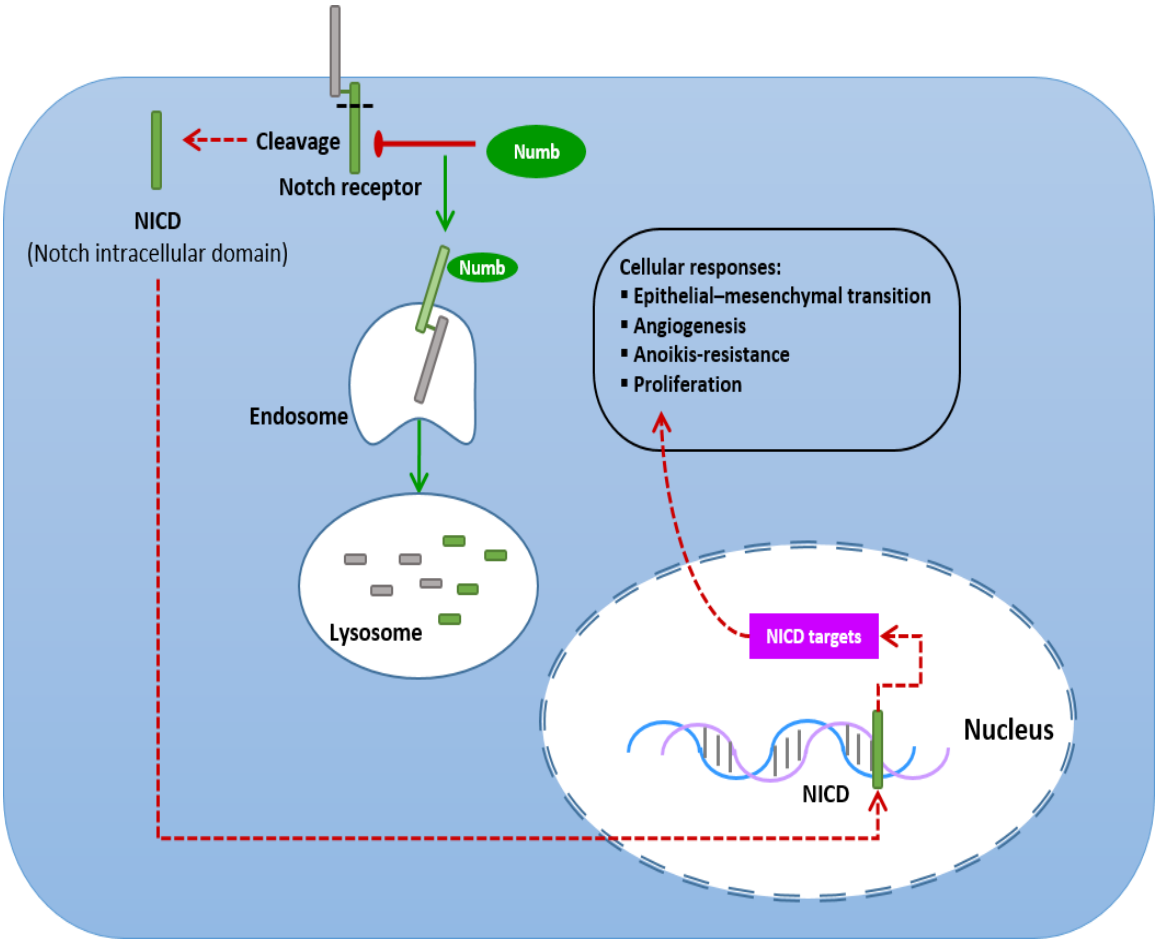


Figure.1.2 NUMB-regulated NOTCH signal transduction

Overview of NUMB mediated inhibition of NOTCH signaling pathways. Upon ligand binding, NOTCH undergoes two cleavages and generates an intracellular fragment NICD, which is subsequently imported into the nucleus and functions as transcription co-activator. NOTCH-regulated genes are widely involved in a number of cellular events. NUMB directly binds to NOTCH at the membrane periphery, which promotes NOTCH internalization and lysosomal degradation.

1.1.2.2 NUMB stabilizes p53 to decrease immortality of cancer cells

Tumor protein p53, also known as TP53 or the “guardian of the genome”, is the first identified tumor suppressor gene, originally described in 1979, and is also the most

commonly mutated gene in all cancer types and most widely studied tumor suppressor in research. At the cellular level, p53 is the central controller of cell mortality and plays a pivotal role in growth arrest, senescence, and apoptosis in response to a broad array of cellular damage events (Levine, 1997).

The human p53 gene encodes a 43.7-kDa (53-kDa on SDS PAGE) protein with several evolutionarily conserved domains including an N-terminal transcription-activation domain, a proline rich domain, a DNA-binding domain and a C-terminal Chomo-oligomerization domain. Typically, p53 maintains a relatively inactive form and low protein level in cells. Upon stimulation of cellular stresses, for example, DNA damage, oxidative stress or osmotic shock, p53 quickly accumulates and switches to an active form through a conformational change, and in turn assembles into a tetramer with full transcription regulation activity. Hundreds of genes are then upregulated by the p53 tetramer, including the key DNA repairing protein PCNA (Proliferating cell nuclear antigen) and cell cycle regulator p21 (CIP1/WAF1, cyclin-dependent kinase inhibitor).

The regulation, and activation of p53 mainly relies on rapid post-translational modifications. Previously, MDM2 (mouse double minute 2 homolog) has emerged as the principal cellular antagonist of p53 (Moll and Petrenko, 2003). MDM2 was first found to be associated with p53 (Momand et al., 1992), and subsequently identified as a E3-ubiquitin ligase that mediates the p53 ubiquitination and lysosomal degradation (Haupt et al., 1997; Kubbutat et al., 1997). p53 and MDM2 form an auto-regulatory feedback loop. MDM2 is a direct transcription target gene of p53, thus active p53 stimulates MDM2 expression. In turn, MDM2 blocks p53 activity by mono-ubiquitinating p53 for degradation in both nuclear and cytoplasmic proteasomes. Cells have also employed various mechanisms to disrupt the balance of this feedback loop in response to stresses. DNA damage can promote the phosphorylation of both p53 and MDM2, thereby preventing the interaction and stabilizing p53 (Lakin and Jackson, 1999). Conversely, receptor tyrosine kinase stimulated Akt (RAC-α serine/threonine-protein kinase) activation phosphorylates MDM2 to promote MDM2 nuclear import, which destabilizes p53 in the nucleus (Zhou et al., 2001).

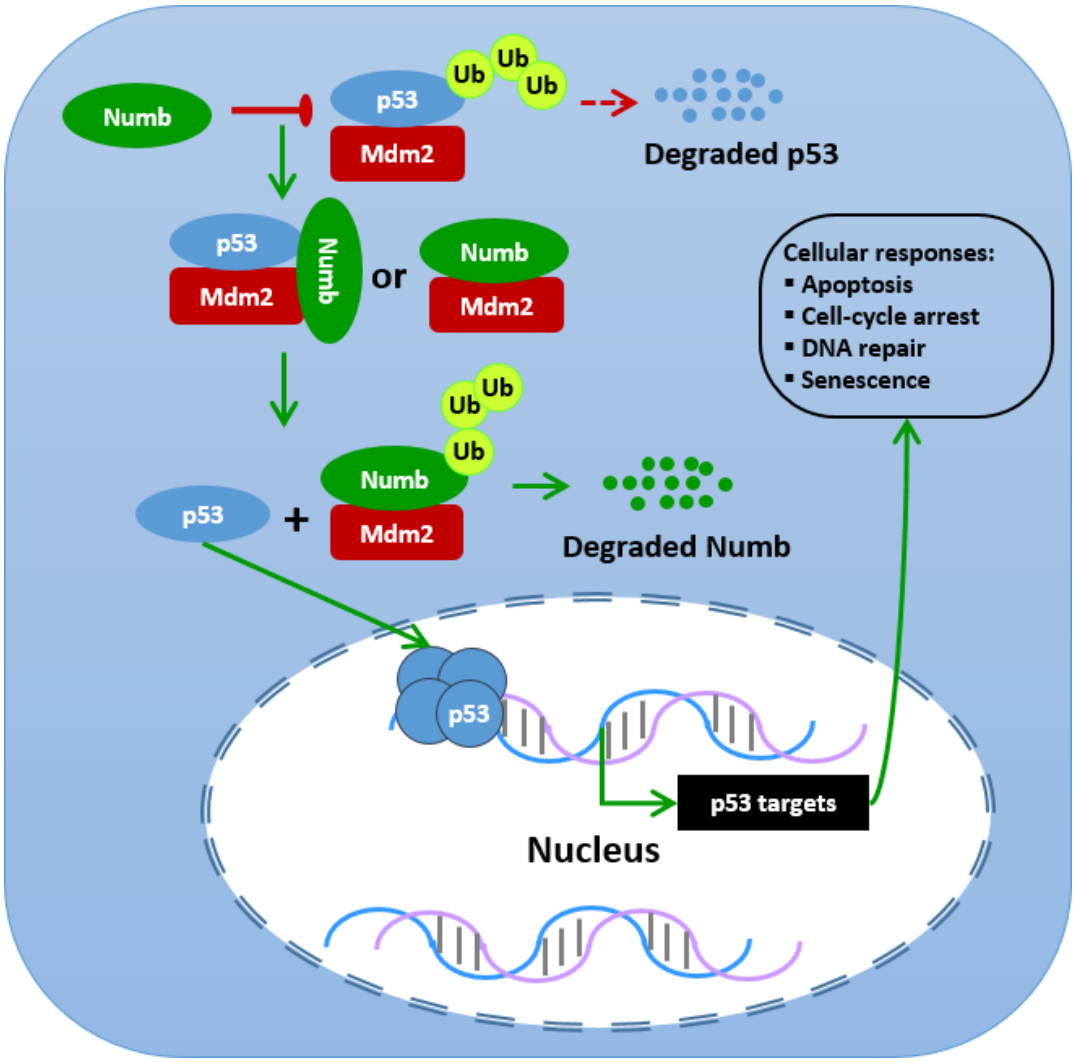


Figure.1.3 Control of p53 activity by NUMB and MDM2

Overview of the regulation of p53 signaling pathway. Typically, p53 maintains a relatively inactive form and low protein level in cells. Upon stimulation, p53 quickly accumulates and switches to an active form by assembling a tetramer that functions as an activator of transcription. p53-regulated genes are widely involved in cell death related events as shown in the figure. MDM2 is an E3 ligase that binds to p53 and promotes p53 degradation. NUMB-p53 binding prevents p53 from MDM2 mediated ubiquitination thus attenuating the immortality of cells.

NUMB has been found to associate with MDM2 (Juven-Gershon et al., 1998) but the function of this interaction was not elucidated until the tricomplex NUMB/MDM2/p53 was described a decade later (Colaluca et al., 2008). In this complex, NUMB binding prevents p53 from ubiquitination by MDM2, thus increasing p53 level in cells (Fig.1.3). In primary breast cancer cells lacking NUMB, p53 level decreases due to ubiqutination which in turn

promotes resistance to chemotherapeutics. In addition, the NUMB-p53 interaction was further found to be regulated by Set8 (N-lysine methyltransferase SETD8) (Dhami et al., 2013). The dynamic methylation of NUMB Lysine-158 and Lysine-163 regulates the interaction between NUMB and p53. Overexpression of Set8 interrupts the NUMB-p53 interaction and abolishes NUMB-induced apoptosis. These findings, combined with the parallel NUMB-NOTCH signaling, predict a potential of adjuvant therapy design for NUMB-defective cancer types that usually display poor prognosis.

1.1.3 The controversial functions of NUMB in cancer

To date, it is known that NUMB has multiple functions at both cellular and tissue levels due to the vast number of NUMB binding partners and the functional diversity among NUMB isoforms. In addition to the negative regulatory effects on NOTCH deregulation and p53 instability, NUMB is also known to directly bind at least three receptor tyrosine kinases (RTKs) (EGFR/EphB2 /TrkB) and attenuate the activation of RTK downstream signaling (Jiang et al., 2012; Nishimura et al., 2006; Zhou et al., 2011). Even though direct evidence is lacking to make a solid conclusion, it is speculated that NUMB may act as a universal regulator for endocytosis of membrane localized proteins. In a similar mechanism for NOTCH, NUMB likely deregulates RTK activity by promoting endocytosis and subsequent degradation of these receptors.

In contrast, NUMB also exhibits some oncogenic properties in an isoform dependent manner. As mentioned above, the four major NUMB isoforms are generated by alternative splicing in exon3 and exon9, and named based on their molecular weights: p72 (PTB^L/PRR^L), p71 (PTB^S/PRR^L), p66 (PTB^L/PRR^S) and p65 (PTB^S/PRR^S). It has been well established that PRR^L and PRR^S isoforms have distinct functions in developmental models of several organs and tissues. PTB^L and PTB^S may also differ in function since p72 and p66 tend to localize to the cell membrane periphery while p71 and p65 generally distribute in the cytoplasm. Consistently, PRR^L isoforms, but not PRR^S isoforms, have been found prominent in a number of cancer types from a large-scale transcript analysis of clinical samples, including breast cancer, colon cancer and lung cancer (Misquitta-Ali et al., 2011). In some lung cancer cells it has also been shown that inclusion of exon 9

correlates with cell proliferation (Bechara et al., 2013; Westhoff et al., 2009). In addition, the RBM5/6 and RBM10 splicing factors, which are responsible for NUMB exon 9 inclusion or exclusion, regulate cancer cell proliferation conversely (Bechara et al., 2013). Together, these observations clearly reveal that NUMB-PRR^{L} isoforms are potentially oncogenic, and these isoforms contribute to the activation of at least one common signaling pathways of cell proliferation in both development and tumorigenesis.

1.2 Receptor tyrosine kinases in cancer

1.2.1 Tyrosine phosphorylation

Phosphorylation is the most common reversible post-translational-modification (PTM) of proteins. 518 protein kinases have been identified in the human genome, constituting about 1.7% of all human genes. These kinases, as well as kinase substrates and phosphatases, are involved in many different cellular events, including metabolism, transcription, cell cycle progression, cytoskeletal rearrangement and cell movement, apoptosis and differentiation (Manning et al., 2002), and deregulation of kinase activity is a hallmark of a multitude of cancer types. To date in cancer treatment, most approved target therapeutics are either small molecule kinase inhibitor or receptor kinase inhibition antibodies.

Serine/threonine kinase activities were first reported in 1954 (Burnett and Kennedy, 1954). Nearly 25 years later, tyrosine phosphorylation was described (Eckhart et al., 1979; Sefton et al., 1980) even though in 1933 phosphotyrosine synthesis already revealed the theoretical possibility of tyrosine phosphorylation (Levene and Schormueller, 1933). The human genome harbors 90 protein tyrosine kinases (PTKs), while most other (~400) kinases are serine/threonine kinases (Manning et al., 2002). Compared to serine/threonine phosphorylation, tyrosine phosphorylation is relatively rare and contributes approximately 5% of all amino acid phosphorylation (Ushiro and Cohen, 1980). Unlike most serine or threonine phosphorylation, tyrosine phosphorylation normally has a very short half-life in cells (seconds in many cases) due to the presence of extremely active phosphotyrosine phosphatases (PTPs) (Hunter, 2014). However, the dynamic nature greatly facilitates RTK role as signal transduction modulators, which rely on fast changes of status. From a structural perspective, tyrosine phosphorylation is also distinct from serine/threonine phosphorylation. The phosphate on a tyrosine residue is linked to the O^4 position of the phenolic ring, which is much farther away from the peptide backbone, compared to the phosphate on the β-OH groups on serine and threonine residues. This provides a more recognizable site for specific binding, thereby allowing the evolution of highly-specific tyrosine kinases, p-Tyr-binding domains and phosphatases (Hunter, 2014).

In animals, a three-part molecular toolkit is employed to regulate tyrosine phosphorylation and p-Tyr-mediated signaling: PTKs, protein tyrosine phosphatases (PTPs) and p-Tyr-recognition domains, respectively, as "writer", "erasers" and "readers" of p-Tyr modifications. PTKs add p-Tyr modifications, PTPs remove the modifications, while p-Tyr-recognition domains readout the modifications. Besides the 90 PTKs, the human genome encodes 107 PTPs or dual specificity phosphatases, and hundreds of p-Tyr-recognition proteins including 121 members of the Src homology 2 (SH2) domain family (Liu et al., 2012; Liu et al., 2006; Liu et al., 2011). Based on all these modules, humans have developed a multi-layered regulation network that precisely tunes the tyrosine phosphorylation dynamics for over 10,000 sites throughout the genome (Liu et al., 2012; Tan et al., 2009).

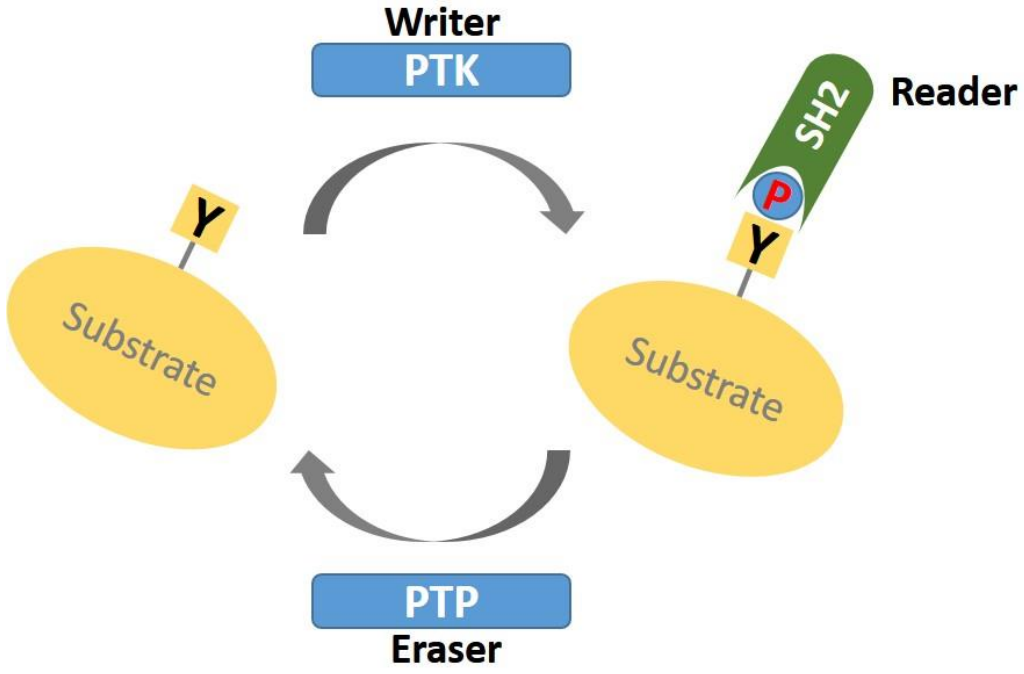


Figure.1.4 Three-part toolkit of tyrosine phosphorylation

Three-part modeling system of p-Tyr signaling. The "writer" PTKs phosphorylate select tyrosine residues in the substrate proteins. The p-Tyr residues are recognized by the "reader" p-Tyr-recognition domain containing proteins (SH2, for example). And the "eraser" PTPs remove the phosphate groups from the tyrosine residues in the substrate proteins.

1.2.2 Tyrosine phosphorylation mediated cell signaling

Tyrosine kinases are generally divided into two major families based on their sub-cellular localizations: the transmembrane receptor tyrosine kinase (RTK) family with 58 members which fall into 20 subfamilies, and the cytoplasmic tyrosine kinase (or non-receptor tyrosine kinase) family with 32 members which falls into at least 9 subfamilies. All RTKs have a similar molecular architecture, containing an amino terminus ligand binding site in the extracellular region, a single transmembrane domain, an intracellular tyrosine kinase domain, and, in many RTKs, a carboxy terminus tail with multiple p-Tyr sites that are crucial for signal transduction. RTKs serve as an entry portal for a large number of extracellular growth factors, and undergo a similar process to activate and transfer the signal into the cytoplasm. In general, the growth factors (ligands) from the extracellular matrix first bind to the ligand binding site on the RTK extracellular region. This ligand-binding mediates a dimerization between two RTK molecules which is the key event in RTK activation. This dimerization then brings two intracellular kinase domains into close proximity, which in turn phosphorylates the tyrosine residues in the activation loop. The activation loop is a short conserved peptide located in the catalytic core of the kinase. In all tyrosine kinases, the activation loop harbors 1~3 tyrosine residues that, in most cases, are the first tyrosine(s) to be phosphorylated during kinase activation. Structural studies have revealed that phosphorylation of the activation loop enables kinase substrate and ATP binding (Hubbard et al., 1994; Lemmon and Schlessinger, 2010). After activation of the kinase domain by intermolecular auto-phosphorylation, the active kinase domain further generates more p-Tyr sites in the intracellular region for recruiting cytoplasmic signaling proteins (p-Tyr-recognition proteins) to initiate the cytoplasmic signaling transduction (Ullrich and Schlessinger, 1990). It should be noted that the growth factors can facilitate the dimerization via at least three different mechanisms: (1) ligand-induced receptor-mediated dimerization, which is elucidated from structural studies of the best known ErbB (epidermal growth factor receptor) RTK subfamily. Ligand binding results in a conformational change in ErbB extracellular regions, which in turn exposes the dimerization sites (Burgess et al., 2003); (2) Dimeric ligands, a number of RTKs, including PDGFR, KIT and their other homologs, lack a dimerization site in the extracellular region

and their dimerization is mediated by the direct interaction between two ligand molecules (Bae et al., 2009); (3) Ligand and accessory molecule-mediated dimerization, which relies on binding of both ligands and other effectors from the extracellular matrix (Schlessinger et al., 2000). Intriguingly, a small subset of RTKs can dimerize even in the absence of activation ligands, meaning a ligand-independent self-activation (Lemmon and Schlessinger, 2010).

Cytoplasmic tyrosine kinases are normally activated via cross-activation with RTKs, in which the kinase domain of an activated RTK directly interacts with a cytoplasmic tyrosine kinase and phosphorylates its activation loop (Lemmon and Schlessinger, 2010). For instance, the cytoplasmic Src kinase specifically localizes at the cell membrane periphery where it can interact with at least two RTKs, ErbB1 (EGFR) and ErbB2. Once activated by the EGF (epidermal growth factor) ligand, ErbB1 and ErbB2 sequentially activate the Src kinase in the cytoplasm (Kim et al., 2005).

Following ligand-stimulated RTK activations and sequential cross-activations of non-receptor tyrosine kinases, active kinases will phosphorylate a number of downstream substrate proteins. Interestingly, profiling of the global tyrosine phosphorylation network following growth factor stimulation exhibited that, in many cases, the highly phosphorylated proteins are the RTKs and non-receptor tyrosine kinases. In addition to the activation loop, tyrosine phosphorylation is also frequently found at other sites in both the kinase and non-catalytic regions. For example, the four members of the ErbB RTK subfamily contain 72 p-Tyr sites (ErbB1/EGFR-22, ErbB2-13, ErbB3-14, ErbB4-23) in total, and the non-receptor tyrosine kinase Src contains at least 9 p-Tyr sites. As shown in Fig.1.5, a number of well-characterized biologically-functional p-Tyr sites have been illustrated in the intracellular region of EGFR receptor. Most of these p-Tyr sites are generated by intermolecular auto-phosphorylation, while at least two sites (p-Tyr-845/p-Tyr-1101) are phosphorylated by Src tyrosine kinase. These p-Tyr sites can be broadly classified into two groups: signaling-related sites and self-regulatory-related sites, which act conversely in the regulation of RTK activity/signaling transduction. For the signaling-related p-Tyr, different sites are recognized by different p-Tyr-recognition (SH2) and phosphotyrosine-binding domain (PTB) proteins, thereby initiating the assembly of

different signaling complexes in the cell membrane periphery to transduce growth factor stimulated signaling. Two major signaling pathways, the PI3K (phosphatidylinositol-4,5-bisphosphate 3-kinase) signaling pathway and the MAPK (mitogen-activated protein kinase) signaling pathway which are also termed as survival pathway and proliferation pathway in cancer biology, are activated by phosphorylation of Y-1068/1086/1148 and Y-1101 respectively. Although not well elucidated, some other p-Tyr sites in the EGFR intracellular region may also promote the activation of PI3K or MAPK signaling pathways, because *in vitro* SH2-p-Tyr binding assay revealed more p-Tyr sites in EGFR binding to PI3K/MAPK-related SH2 proteins such as Grb2. The other group of p-Tyr sites, which also play a role as docking sites for SH2 proteins, are mainly involved in negative self-regulation of EGFR. p-Tyr-1045 and p-Tyr-1173 are two key self-regulatory sites distinct in their regulation mechanisms. p-Tyr-1045 is recognized by a SH2-containing E3 ubiquitin ligase Cbl. Upon Cbl binding, the nearby lysine residues are ubiqutinated which serve as a sorting signal for targeting activated EGFR to coated pits for endocytosis and subsequent lysosomal degradation (Ettenberg et al., 1999; Grovdal et al., 2004; Visser Smit et al., 2009). p-Tyr-1173 binds to the SH2 domain-containing protein tyrosine phosphatase (SHP-1), which dephosphorylates the EGFR intracellular region to attenuate signaling (Keilhack et al., 1998). It is now well known that tyrosine phosphorylation can further introduce other types of post-translational-modification to RTKs via recruiting respective SH2 containing PTM modulators, which, in many cases, conversely regulate the activities of RTKs in a feedback regulatory loop (Hunter, 2007; Lemmon and Schlessinger, 2010).

Because p-Tyr-SH2 binding only depends on a small peptide fragment around the phosphotyrosine residue, these SH2-containing proteins can potentially act universally. For instance, the Cbl recognizes a number of active RTKs, including but not limited to KIT, FLT1, FGFR1, FGFR2, PDGFRA, PGGFRB, EGFR, CSF1R, EPHA8 and KDR, as summarized in the Uniprot Database. Since RTKs stimulate signaling by employing the same set of SH2-containing proteins, RTKs are highly functionally-redundant especially in the activation of PI3K and MAPK signaling pathways, which contribute prominently to cancer cell immortality and fast proliferation. Therefore, targeting tyrosine kinase has become a promising strategy in treating specific cancer types in the clinic. As of February 2016, 46 target therapeutics have been clinically approved by FDA for cancer treatment,

of which 28 are small molecule kinase inhibitors and 18 are humanized antibodies. It is easy to tell the significance of tyrosine kinase inhibitors (TKIs) simply by the statistics: 25 of 28 small molecule inhibitors mimic the ATP structure to block the catalytic core of tyrosine kinases, while 6 of 18 antibodies are designed to interrupt ligand stimulated RTK dimerization. In clinical diagnosis, determination of PTK gene amplification or protein expression level has also become standard procedure for a number of cancer types. For instance, approximately 30% of breast cancers are diagnosed as having ErbB2 gene amplification (HER2 positive).

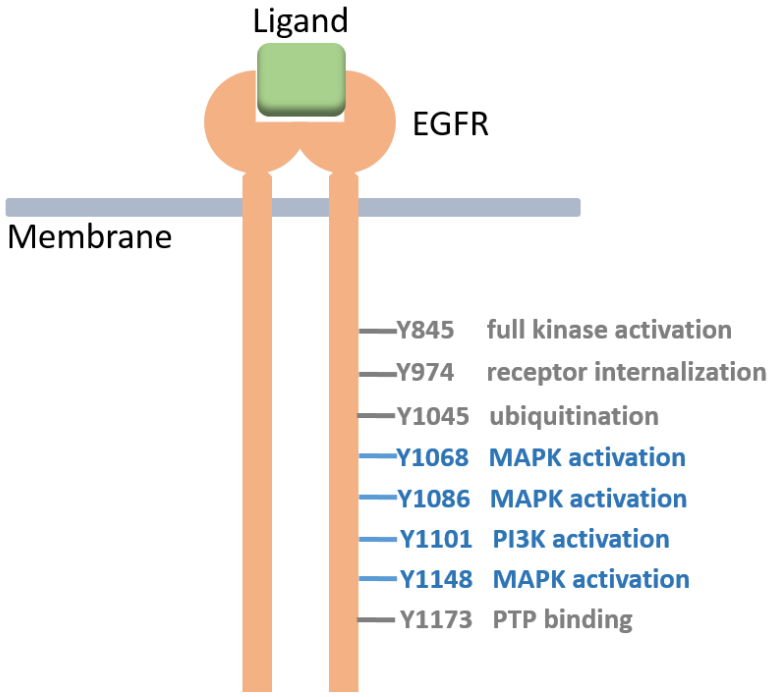


Figure.1.5 Functions of select p-Tyr sites in the EGFR receptor

Multiple p-Tyr sites are generated in the intracellular region of EGFR following ligand stimulated EGFR dimerization and activation. Most sites are produced by intermolecular auto-phosphorylation between dimerized EGFR molecules. At least two sites, p-Tyr-845 and p-Tyr-1101 are phosphorylated by Src tyrosine kinase that associates with EGFR at the cell membrane periphery. The p-Tyr sites serve as docking sites for different SH2-containing proteins, which plays different roles after binding to EGFR. Sites marked in grey are self-regulation sites: (1) p-Tyr-845 stabilizes the kinase in an active status; (2) p-Tyr-974 promotes EGFR internalization; (3) p-Tyr-1045 binds to an E3 ubiquitin ligase Cbl thus increasing EGFR ubiquitination and degradation; (4) p-Tyr-1173 binds to PTP that attenuates EGFR phosphorylation level. Sites marked in blue are signaling transduction sites. p-Tyr-1068/1086/1148 all stimulate the MAPK signal pathways while p-Tyr-1101 only stimulates the PI3K signaling pathway.

Despite the quickly-expanding understanding of RTK-mediated signal transduction, our knowledge of the spatiotemporal dynamics of how RTK signaling controls specific cellular events has lagged behind (Volinsky and Kholodenko, 2013). Two great challenges have been encountered along the way. One limitation is that a reliable high-throughput approach to quickly profile the kinase-substrate pairings in the kinome is lacking. Even though not as abundant as serine/threonine phosphorylation in the human proteome, over 10,000 p-Tyr sites have been identified in human due to recent improvements in p-Tyr-enrichment and mass-spectrometry (MS) approaches, while most are non-PTK p-Tyr-sites. Many non-PTK p-Tyr sites have been proven functional and implicated in a wide range of cellular events (Reynolds et al., 2014). The best-studied Src tyrosine kinase is the first described oncogene and PTK (Hunter and Sefton, 1980), which has only 67 substrates (PhosphoSitePlus Database) that have been verified *in vitro* and *in vivo* almost four decades after its discovery. Many p-Tyr sites in these substrates have been shown to be correlated with the transformation ability of Src kinase (Reynolds et al., 2014). The slowly emerging kinase-substrate pairings are largely due to the promiscuous nature of PTKs, which makes the conventional kinase-assay-based approaches not necessarily reliable. In addition, the functional redundancy and cross-activation of PTKs also limits the application of MS based *in vivo* approaches.

Even if the kinase-substrate is systematically paired, we are also facing another challenge in deciphering the complex signaling network. A good example exhibiting the complexity of RTK signaling is from studies on PC12 cells. Treatment by two different growth factors, NGF (nerve growth factor) and EGF (epidermal growth factor), resulted in different biological effects in PC12 cells (Morooka and Nishida, 1998). NGF initiates the differentiation of PC12 cells while EGF only stimulates cell proliferation. However, this is not caused by the activation of different RTKs that NGF/EGF target (Tyk/ErbB, respectively), but because of the different durations of the active status of MAPK signaling pathway, which was confirmed in another study using a light-controlled MAPK activation system (Zhang et al., 2014). This interesting observation has raised many questions yet to be answered, but clearly demonstrates the complex nature of the p-Tyr-mediated signaling network. If simply focusing on the network of ErbB subfamily, some progress has been made to draw a relatively comprehensive network of regulation map, as reviewed by

Yarden and Sliwkowski (Yarden and Sliwkowski, 2001). The overall signaling network is divided into three major layers: input, signal processing and output. (1) The input layer consists of four ErbB subfamily members and a number of extracellular growth factors (ligands). ErbB receptors can form either heterodimers or homodimers, and different growth factors favor different dimers. These diversified dimers among ErbB receptors consequently generate different combinations of p-Tyr sites (SH2 docking sites) in their intracellular regions, thereby recruiting different groups of SH2-containing signaling complexes. At this input layer, the relative abundance of growth factors and ErbB proteins differentiate the signal output. (2) In the signal-processing layer, as expected, the assembly of different SH2-containing signaling complexes in the membrane periphery will result in the activation of different signaling processing components. Similarly, the abundance of these SH2-containing proteins will determine the direction of signal transduction. In addition, several feedback loops exist between the input layer and the signal-processing layer, for example, Cbl E3 ligase binds to p-Tyr sites in ErbB and the subsequent ubiquitination promotes ErbB degradation, as introduced earlier. (3) The output layer is the least-understood and most complicated part in this network. Because most pTyr sites identified in Src kinase substrates contribute to Src transformation ability (Reynolds et al., 2014), it can be speculated that a large number of uncharacterized tyrosine phosphorylated proteins may be implicated in a broad range of cellular events in tumorigenesis.

1.2.3 Endocytosis regulates RTK signaling

Endocytosis is a process in which cells take in extracellular materials by engulfing and fusing them with the cell membrane. Depending on the materials cells take up, endocytosis is further classified as pinocytosis (cell drinking) and phagocytosis (cell eating). A common destination of the internalized material is digestion/degradation in the lysosome, which is why endocytosis is termed as cell drinking and eating. On the other hand, cells also evacuate intracellular materials via an opposing process called exocytosis. Endocytosis and exocytosis are key cellular events and essential for material exchange between cells and the environment. Another well-known function of endocytosis is found in the immune system, in which some immune cells actively capture and internalize pathogens, foreign particles or debris from dying cells by phagocytosis.

Besides the material exchange, endocytosis is one of the major regulators of the activity and signaling of RTKs, as well as many other receptors on the cell membrane. These transmembrane proteins are internalized along with a portion of the cell membrane, inserted in the lipid bilayer of endosomal vesicles and subsequently degraded in the lysosome. At the cell surface, endocytosis is constitutively occurring for internalizing RTKs, which is crucial for RTK function in both an inactive status and a ligand-stimulated active status of cells.

In the inactive cells, endocytosis is a process required for the renewal of RTKs on the cell membrane. Internalized RTKs have two distinct destinations following endocytosis, recycling back to the membrane from early endosomal vesicles or compartmentalizing in late endosomal vesicles. This process is also termed endosomal sorting and is regulated by a number of effectors that dynamically control the balance between recycling and degradation. The rate of RTK endocytosis is usually slower compared to that of RTK synthesis and delivery from Golgi apparatus to membrane. Thus, RTKs can retain a relatively high abundance on the cell membrane, which is thought to facilitate an expeditious response to extracellular growth signals (binding of ligands) (Goh and Sorkin, 2013). Altogether, endocytosis, protein synthesis and delivery and the preference of endosomal sorting decide the renewal rate of RTKs and assist in maintaining RTK stability on the cell membrane.

Upon ligand stimulated activation, cells also employ endocytosis to regulate the signaling from RTKs. As introduced earlier, ligand binding quickly triggers auto-phosphorylation in the intracellular regions of RTKs to generate multiple p-Tyr sites, which then recruit SH2 containing proteins that function in signaling transduction or RTK self-regulation (Fig.1.5). During this process, some endocytic proteins can be recruited to the RTKs, thus initiating a fast endocytosis. Conventionally, endocytosis is thought to negatively regulate RTK activity and signaling by decreasing RTK abundance on the cell membrane and promoting RTK degradation, even though a portion of internalized RTK is recycled back to the membrane. Studies on the endocytosis of EGFR have provided solid evidence to support this model. Upon activation, the EGF-EGFR complex is detectable in endosomal vesicles within 2~5 minutes following EGF stimulation (Beguinot et al., 1984; Haigler et al., 1979).

Simultaneously, the free (un-liganded) EGFR still undergoes endocytosis but at a rate at least 10-fold slower (Wiley et al., 1991). Following internalization, EGFR is degraded in the lysosome, thus overall EGFR activity and downstream MAPK and PI3K signaling are attenuated.

This model has been challenged by some recent studies, which reveal that ligand-stimulated RTK endocytosis may play a role as both a positive and negative regulator of RTKs. RTKs are still able to signal in the endosomal system before degradation and the effect of endocytosis varies under different circumstances (Kermorgant and Parker, 2008; Kermorgant et al., 2004; McPherson et al., 2001; Miaczynska et al., 2004). It is unsurprising that membrane-localized EGFR is confirmed as the major contributor to the activation of MAPK and PI3K signaling. However, EGF bound EGFR is found to remain active in endosomal vesicles, while the EGFR intracellular region is still exposed to the cytoplasm and able to recruit SH2 containing proteins for assembly of signaling complexes (Brankatschk et al., 2012; Sousa et al., 2012). It is estimated that endosomal EGF-EGFR may contribute ~10% to the activation of MAPK and PI3K signaling in specific environments. In addition, it is reported that endocytosis of EGFR is required for the activation of a subset of signaling transducers, suggesting endosomal EGF-EGFR may primarily activate some side signaling pathways in the cytoplasm (Vieira et al., 1996). The most promising evidence comes from a study of MET (hepatocyte growth factor receptor) RTK, which clearly demonstrates the pivotal role of endocytosis in MET oncogenicity (Joffre et al., 2011). This study takes advantage of a MET mutant that exhibits accumulation on the endosome due to increased endocytosis/recycling activity and decreased levels of degradation. Blocking endocytosis has little effect on MET auto-phosphorylation but significantly inhibits anchorage-independent growth, *in vivo* tumorigenesis and *in vivo* metastasis. In summary, RTK function results not only from ligand stimulated auto-phosphorylation but also from the altered subcellular localizations during endocytosis. Endocytosis may bring an extra spatiotemporal regulatory layer to the RTK signaling network.

Endocytosis has also been defined as a potential therapeutic target for some specific cancer types. A study in EGFR-positive lung cancer reports that abnormal endocytosis is

associated with intrinsic resistance to the EGFR small molecule inhibitor gefitinib. The deregulation of Rab25, an endocytic protein, results in drug resistance by inducing continuous EGFR endocytosis, and the knock-down of Rab25 is able to re-sensitize these resistant cells to gefitinib (Jo et al., 2014). Drug resistance to target therapeutics is a major problem facing current cancer research, and is common in all RTK positive cancer types (Holohan et al., 2013). The drug resistance can be intrinsic, which is present before treatment, or acquired during treatment. For instance, although diagnosed as HER2 positive, around 70% of patients have little or no response to the ErbB2 inhibitor Herceptin (Baselga et al., 1996). In the remaining 30% of patients, many of them will develop acquired resistance quickly after receiving the treatment, usually within months, which is caused by therapy-induced adaptive responses or positive selection of an intrinsically-resistant subpopulation. The mechanism of resistance is highly complicated and yet not systematically evaluated in large scale studies. From a number of individual reports, there are two common causes: (1) mutations in the ATP binding site prevent binding of the small molecule kinase inhibitors (Coco et al., 2012; Kobayashi et al., 2005); (2) activation of alternative signaling pathways bypasses the drug inhibition (Lu et al., 2001). This study first reveals that abnormal endocytosis contributes to drug resistance in RTK positive cancer types. However, the mechanism is not clearly elucidated, and it is only speculated that continuous EGFR endocytosis induced by Rab25 overexpression stimulates gefitinib-resistance related signaling.

1.2.4 ALK-positive cancers and ALK target therapeutics

The receptor tyrosine kinase Anaplastic lymphoma kinase (ALK), also known as CD246 (cluster of differentiation 246), was first described during the identification of a 2;5 chromosomal translocation in anaplastic large-cell non-Hodgkin's lymphoma (ALCL) (Morris et al., 1994; Shiota et al., 1994). ALK belongs to a small RTK subfamily, consisting of only two members ALK and LTK (leukocyte receptor tyrosine kinase), as defined by their unique extracellular domains containing a glycine-rich region (Palmer et al., 2009). Based on their kinase domain similarities, ALK/LTK are closely related to the insulin receptor tyrosine kinase subfamily. Previously, a few potential ALK ligands were reported but none have been well characterized (Murray et al., 2015; Stoica et al., 2001;

Stoica et al., 2002). Compared to many other well-studied RTKs, ALK is not functionally significant in physiology, and only highly-expressed in the neural system during the early developmental stage (Morris et al., 1994).

ALK did not draw attention from researchers until multiple reports in 2008 defined it as a biomarker and therapeutic target in neuroblastomas (Carén et al., 2008; Chen et al., 2008; George et al., 2008; Janoueix-Lerosey et al., 2008; Mossé et al., 2008), including four *Nature* reports. Since then, ALK has been shown to be involved in the initiation and progression of many different cancer types, including lymphomas, neuroblastoma, non-small-cell lung cancer (NSCLC), colorectal cancer and melanomas (Hallberg and Palmer, 2013). The activity of ALK can be deregulated by mutation, gene amplification, translocation fusion and as reported recently, alternative transcription initiation (Wiesner et al., 2015). It should be noted that the ALK locus is a hotspot of translocation, though the reason remains unclear. Over 30 ALK kinase domain containing fusion proteins with at least 22 fusion partners have been identified in different cancers (Fig.1.6). The frequencies of these translocations vary, except for the EML4 (echinoderm microtubule-associated protein-like 4) fusion which has been diagnosed in 4%~7% of NSCLC and shown to promote and maintain the malignant behavior of the cancer cells (Sasaki et al., 2010; Soda et al., 2007). The oncogenic fusion partners share characteristics, which generally promote the dimerization of the ALK kinase domain, a mechanism that mimics the ligand stimulated dimerization of RTKs in the cytoplasm. Similar to other RTK positive tumors, deregulated ALK promotes tumorigenesis by constitutively stimulating MAPK and PI3K signaling pathways. Because ALK normally maintains a low level and is likely not functionally essential in adults, ALK inhibition exhibited remarkably strong drug tolerance, meaning a higher dose is applicable in clinic with less adverse effects (Crescenzo and Inghirami, 2015).

The first generation ALK small molecule inhibitor crizotinib (Xalkori®, Pfizer) was granted an accelerated approval by FDA in 2011 for the treatment of advanced NSCLC with EML4-ALK fusion, which set a record from discovery to therapy progressing in less than three years. In addition, the second generation inhibitor alectinib (Alecensa®, Roche) was also granted an accelerated approval in December, 2015.

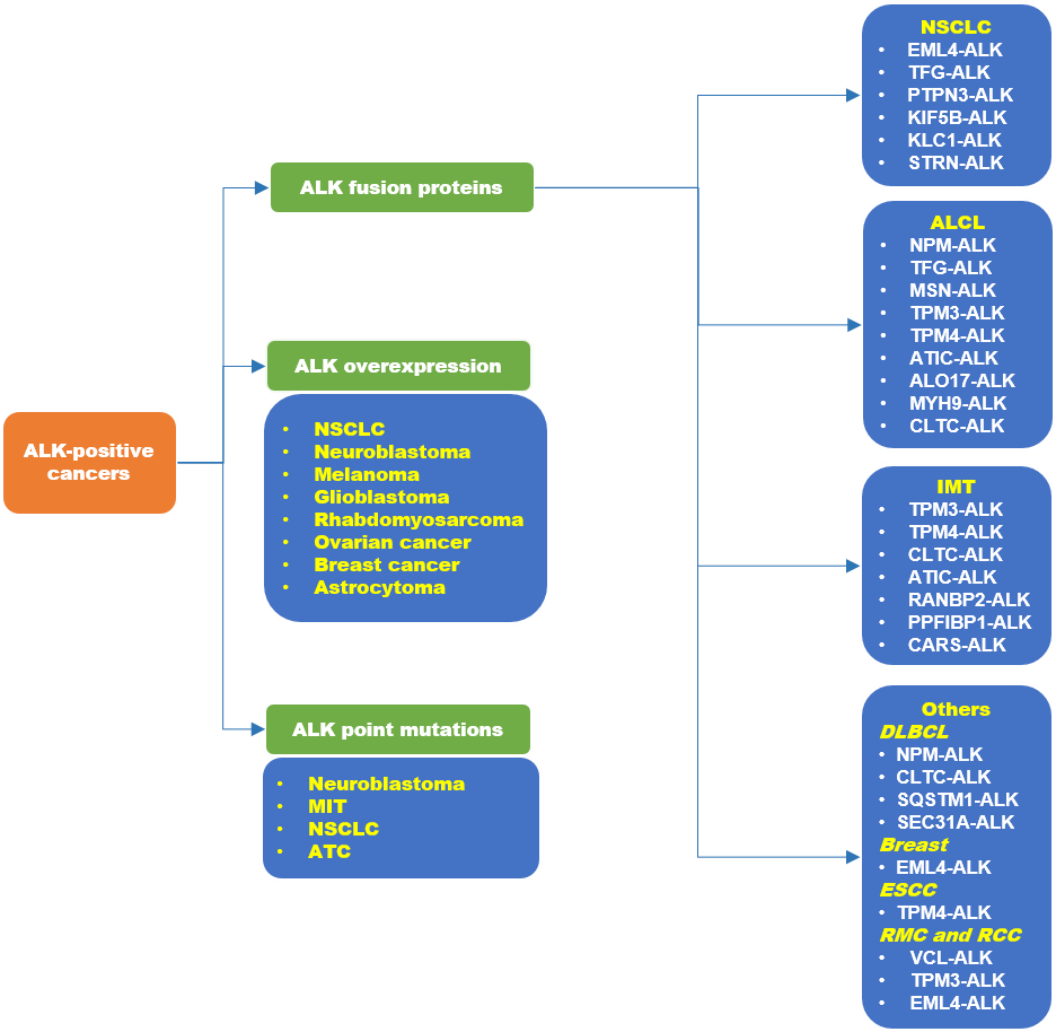


Figure.1.6 Oncogenic forms of ALK

Overview of the ALK de-regulated forms in tumorigenesis. Similar to that in other RTK positive cancer types, overexpression and constitutively active (kinase) mutations are common in ALK positive cancers. It's intriguing that ALK forms over 30 different translocation fusion proteins with at least 22 fusion partners in a multitude of cancer types. Adapted from (Hallberg and Palmer, 2013).

Despite the oncogenic potential and therapeutic value of ALK, the mechanism of ALK regulation remains elusive. Currently most efforts have been put toward the investigation of ALK deregulation in different cancer types or the mechanism of crizotinib resistance. In the following chapters, we will present a mechanistic insight into NUMB mediated ALK regulation as a potentially universal negative regulatory system for RTKs.

1.3 Scope of thesis project

NUMB is a multifaceted protein that is involved in a broad range of cellular events. In cancer biology, NUMB has exhibited opposing biological effects, either as a tumor suppressor via attenuating oncogenic NOTCH signaling, assisting in p53 stabilization, or as a potential oncogene. The oncogenic nature of NUMB is not mechanistically elucidated but supported by a number of clinical sample based studies. At the molecular level, the complex nature of NUMB function is speculated to be caused by the heterogeneity of NUMB interacting proteins in different genetic backgrounds.

To uncover the mechanisms of NUMB functions, particularly in cancer biology, we designed a stepwise strategy to first decipher NUMB binding specificity and identify novel NUMB binding partners via an integrated high-throughput approach. It is intriguing that a number of RTKs bind NUMB both in *vitro* and *in vivo*, revealing that NUMB may act as a universal regulator of RTKs. These findings, once further validated, could potentially contribute to the personal diagnosis and therapy of RTK positive cancer types, for example the combination of RTK deregulation and NUMB depletion may indicate a more severe prognosis. Our following work mainly focuses on the ALK receptor, which is the therapeutic target in a number of cancers. To date two generations of ALK inhibitors have been clinically approved. We found that NUMB regulates ALK endocytosis, similar to its regulation of NOTCH receptor. Interestingly, following internalization, the two NUMB isoforms p72 and p66 direct ALK into distinct destinations in endosomal sorting. The p72 isoform mainly recycles ALK back to the membrane to sustain the active status of ALK signaling while the p66 isoform attenuates ALK signaling by mediating ALK degradation in the lysosome. Consistently, the two isoforms exhibited opposing effects in anchorage-dependent cancer cell growth. This NUMB-ALK interaction does not rely on ALK phosphorylation, and it is still uncertain whether this interaction responds to ALK activation thus forming a feedback regulatory loop. The promotion of endocytosis has been coupled with RTK activation as reported in many preliminary studies, and the endocytosis is thought to be a significant strategy employed by the cells to self-balance RTK activity.

Overall, our findings provide a mechanistic insight into ALK regulation and the tumorigenic potential of the p72-NUMB isoform, and further reveal markers of potential clinical value.

1.4 Reference

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

Bae, J. H., Lew, E. D., Yuzawa, S., Tomé, F., Lax, I., and Schlessinger, J. (2009). The Selectivity of Receptor Tyrosine Kinase Signaling Is Controlled by a Secondary SH2 Domain Binding Site. Cell *138*, 514-524.

Bani-Yaghoub, M., Kubu, C. J., Cowling, R., Rochira, J., Nikopoulos, G. N., Bellum, S., and Verdi, J. M. (2007). A switch in numb isoforms is a critical step in cortical development. Developmental Dynamics *236*, 696-705.

Barth, J. M. I., and Köhler, K. (2014). How to take autophagy and endocytosis up a notch. BioMed Research International *2014*.

Baselga, J., Tripathy, D., Mendelsohn, J., Baughman, S., Benz, C. C., Dantis, L., Sklarin, N. T., Seidman, A. D., Hudis, C. A., Moore, J., *et al.* (1996). Phase II study of weekly intravenous recombinant humanized anti-p185HER2 monoclonal antibody in patients with HER2/neu-overexpressing metastatic breast cancer. J Clin Oncol *14*, 737-744.

Bechara, E. G., Sebestyen, E., Bernardis, I., Eyras, E., and Valcarcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Molecular cell *52*, 720-733.

Beguinot, L., Lyall, R. M., Willingham, M. C., and Pastan, I. (1984). Down-regulation of the epidermal growth factor receptor in KB cells is due to receptor internalization and subsequent degradation in lysosomes. Proceedings of the National Academy of Sciences of the United States of America *81*, 2384-2388.

Berdnik, D., Torok, T., Gonzalez-Gaitan, M., and Knoblich, J. A. (2002). The endocytic protein alpha-Adaptin is required for numb-mediated asymmetric cell division in Drosophila. Developmental cell *3*, 221-231.

Blaumueller, C. M., Qi, H., Zagouras, P., and Artavanis-Tsakonas, S. (1997). Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane. Cell *90*, 281-291.

Bork, P., and Margolis, B. (1995). A phosphotyrosine interaction domain. Cell *80*, 693-694.

Brankatschk, B., Wichert, S. P., Johnson, S. D., Schaad, O., Rossner, M. J., and Gruenberg, J. (2012). Regulation of the EGF transcriptional response by endocytic sorting. Science Signaling *5*.

Bray, S. J. (2006). Notch signalling: A simple pathway becomes complex. Nature Reviews Molecular Cell Biology *7*, 678-689.

Brou, C. (2009). Intracellular trafficking of Notch receptors and ligands. Experimental Cell Research *315*, 1549-1555.

Brou, C., Logeat, F., Gupta, N., Bessia, C., LeBail, O., Doedens, J. R., Cumano, A., Roux, P., Black, R. A., and Israël, A. (2000). A novel proteolytic cleavage involved in Notch signaling: The role of the disintegrin-metalloprotease TACE. Molecular cell *5*, 207-216.

Burgess, A. W., Cho, H. S., Eigenbrot, C., Ferguson, K. M., Garrett, T. P. J., Leahy, D. J., Lemmon, M. A., Sliwkowski, M. X., Ward, C. W., and Yokoyama, S. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Molecular cell *12*, 541-552.

Burnett, G., and Kennedy, E. P. (1954). The enzymatic phosphorylation of proteins. The Journal of biological chemistry *211*, 969-980.

Capobianco, A. J., Zagouras, P., Blaumueller, C. M., Artavanis-Tsakonas, S., and Bishop, J. M. (1997). Neoplastic transformation by truncated alleles of human NOTCH1/TAN1 and NOTCH2. Molecular and Cellular Biology *17*, 6265-6273.

Carén, H., Abel, F., Kogner, P., and Martinsson, T. (2008). High incidence of DNA mutations and gene amplifications of the ALK gene in advanced sporadic neuroblastoma tumours. Biochemical Journal *416*, 153-159.

Chen, Y., Takita, J., Choi, Y. L., Kato, M., Ohira, M., Sanada, M., Wang, L., Soda, M., Kikuchi, A., Igarashi, T., *et al.* (2008). Oncogenic mutations of ALK kinase in neuroblastoma. Nature *455*, 971-974.

Chiba, S. (2006). Notch signaling in stem cell systems. Stem cells *24*, 2437-2447.

Coco, S., De Mariano, M., Valdora, F., Servidei, T., Ridola, V., Andolfo, I., Oberthuer, A., Tonini, G. P., and Longo, L. (2012). Identification of ALK germline mutation (3605delG) in pediatric anaplastic medulloblastoma. Journal of Human Genetics *57*, 682-684.

Colaluca, I. N., Tosoni, D., Nuciforo, P., Senic-Matuglia, F., Galimberti, V., Viale, G., Pece, S., and Di Fiore, P. P. (2008). NUMB controls p53 tumour suppressor activity. Nature *451*, 76-80.

Crescenzo, R., and Inghirami, G. (2015). Anaplastic lymphoma kinase inhibitors. Curr Opin Pharmacol *23*, 39-44.

Davis, R. L., and Turner, D. L. (2001). Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. Oncogene *20*, 8342-8357.

Dhami, G. K., Liu, H., Galka, M., Voss, C., Wei, R., Muranko, K., Kaneko, T., Cregan, S. P., Li, L., and Li, S. S. (2013). Dynamic methylation of Numb by Set8 regulates its binding to p53 and apoptosis. Molecular cell *50*, 565-576.

Dho, S. E., French, M. B., Woods, S. A., and McGlade, C. J. (1999). Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. J Biol Chem *274*, 33097-33104.

Dho, S. E., Jacob, S., Wolting, C. D., French, M. B., Rohrschneider, L. R., and McGlade, C. J. (1998). The mammalian numb phosphotyrosine-binding domain. Characterization of binding specificity and identification of a novel PDZ domain-containing numb binding protein, LNX. Journal of Biological Chemistry *273*, 9179-9187.

Dho, S. E., Trejo, J., Siderovski, D. P., and McGlade, C. J. (2006). Dynamic regulation of mammalian numb by G protein-coupled receptors and protein kinase C activation: Structural determinants of numb association with the cortical membrane. Molecular biology of the cell *17*, 4142-4155.

Dooley, C. M., James, J., McGlade, C. J., and Ahmad, I. (2003). Involvement of numb in vertebrate retinal development: Evidence for multiple roles of numb in neural differentiation and maturation. Journal of Neurobiology *54*, 313-325.

Eckhart, W., Hutchinson, M. A., and Hunter, T. (1979). An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. Cell *18*, 925-933.

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

Espinoza, I., and Miele, L. (2013). Notch inhibitors for cancer treatment. Pharmacology and Therapeutics *139*, 95-110.

Ettenberg, S. A., Keane, M. M., Nau, M. M., Frankel, M., Wang, L. M., Pierce, J. H., and Lipkowitz, S. (1999). cbl-b inhibits epidermal growth factor receptor signaling. Oncogene *18*, 1855-1866.

Flores, A. N., McDermott, N., Meunier, A., and Marignol, L. (2014). NUMB inhibition of NOTCH signalling as a therapeutic target in prostate cancer. Nature Reviews Urology *11*, 499-507.

Frise, E., Knoblich, J. A., Younger-Shepherd, S., Jan, L. Y., and Jan, Y. N. (1996). The Drosophila Numb protein inhibits signaling of the Notch receptor during cell-cell interaction in sensory organ lineage. Proc Natl Acad Sci U S A *93*, 11925-11932.

George, R. E., Sanda, T., Hanna, M., Frohling, S., Luther, W., 2nd, Zhang, J., Ahn, Y., Zhou, W., London, W. B., McGrady, P., *et al.* (2008). Activating mutations in ALK provide a therapeutic target in neuroblastoma. Nature *455*, 975-978.

Gho, M., Lecourtois, M., Geraud, G., Posakony, J. W., and Schweisguth, F. (1996). Subcellular localization of Suppressor of Hairless in Drosophila sense organ cells during Notch signalling. Development *122*, 1673-1682.

Goh, L. K., and Sorkin, A. (2013). Endocytosis of receptor tyrosine kinases. Cold Spring Harb Perspect Biol *5*, a017459.

Greenwald, I. (1994). Structure/function studies of lin-12/Notch proteins. Current Opinion in Genetics and Development *4*, 556-562.

Grovdal, L. M., Stang, E., Sorkin, A., and Madshus, I. H. (2004). Direct interaction of Cbl with pTyr 1045 of the EGF receptor (EGFR) is required to sort the EGFR to lysosomes for degradation. Experimental Cell Research *300*, 388-395.

Gulino, A., Di Marcotullio, L., and Screpanti, I. (2010). The multiple functions of Numb. Exp Cell Res *316*, 900-906.

Guo, M., Jan, L. Y., and Jan, Y. N. (1996). Control of daughter cell fates during asymmetric division: Interaction of Numb and Notch. Neuron *17*, 27-41.

Haigler, H. T., McKanna, J. A., and Cohen, S. (1979). Rapid stimulation of pinocytosis in human carcinoma cells A-431 by epidermal growth factor. Journal of Cell Biology *83*, 82-90.

Hallberg, B., and Palmer, R. H. (2013). Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. Nature reviews Cancer *13*, 685-700.

Haupt, Y., Maya, R., Kazaz, A., and Oren, M. (1997). Mdm2 promotes the rapid degradation of p53. Nature *387*, 296-299.

Holohan, C., Van Schaeybroeck, S., Longley, D. B., and Johnston, P. G. (2013). Cancer drug resistance: an evolving paradigm. Nature reviews Cancer *13*, 714-726.

Hubbard, S. R., Wei, L., Ellis, L., and Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature *372*, 746-754.

Hunter, T. (2007). The Age of Crosstalk: Phosphorylation, Ubiquitination, and Beyond. Molecular cell *28*, 730-738.

Hunter, T. (2014). The genesis of tyrosine phosphorylation. Cold Spring Harb Perspect Biol *6*, a020644.

Hunter, T., and Sefton, B. M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. Proceedings of the National Academy of Sciences of the United States of America *77*, 1311-1315.

Iso, T., Kedes, L., and Hamamori, Y. (2003). HES and HERP families: Multiple effectors of the Notch signaling pathway. Journal of Cellular Physiology *194*, 237-255.

Janoueix-Lerosey, I., Lequin, D., Brugières, L., Ribeiro, A., De Pontual, L., Combaret, V., Raynal, V., Puisieux, A., Schleiermacher, G., Pierron, G., *et al.* (2008). Somatic and

germline activating mutations of the ALK kinase receptor in neuroblastoma. Nature *455*, 967-970.

Jiang, X., Xing, H., Kim, T. M., Jung, Y., Huang, W., Yang, H. W., Song, S., Park, P. J., Carroll, R. S., and Johnson, M. D. (2012). Numb regulates glioma stem cell fate and growth by altering epidermal growth factor receptor and Skp1-Cullin-F-box ubiquitin ligase activity. Stem cells *30*, 1313-1326.

Jo, U., Park, K. H., Whang, Y. M., Sung, J. S., Won, N. H., Park, J. K., and Kim, Y. H. (2014). EGFR endocytosis is a novel therapeutic target in lung cancer with wild-type EGFR. Oncotarget *5*, 1265-1278.

Joffre, C., Barrow, R., Menard, L., Calleja, V., Hart, I. R., and Kermorgant, S. (2011). A direct role for Met endocytosis in tumorigenesis. Nat Cell Biol *13*, 827-837.

Juven-Gershon, T., Shifman, O., Unger, T., Elkeles, A., Haupt, Y., and Oren, M. (1998). The Mdm2 oncoprotein interacts with the cell fate regulator Numb. Molecular and cellular biology *18*, 3974.

Keilhack, H., Tenev, T., Nyakatura, E., Godovac-Zimmermann, J., Nielsen, L., Seedorf, K., and Bohmer, F. D. (1998). Phosphotyrosine 1173 mediates binding of the protein-tyrosine phosphatase SHP-1 to the epidermal growth factor receptor and attenuation of receptor signaling. J Biol Chem *273*, 24839-24846.

Kermorgant, S., and Parker, P. J. (2008). Receptor trafficking controls weak signal delivery: A strategy used by c-Met for STAT3 nuclear accumulation. Journal of Cell Biology *182*, 855-863.

Kermorgant, S., Zicha, D., and Parker, P. J. (2004). PKC controls HGF-dependent c-Met traffic, signalling and cell migration. EMBO Journal *23*, 3721-3734.

Kim, H., Chan, R., Dankort, D. L., Zuo, D., Najoukas, M., Park, M., and Muller, W. J. (2005). The c-Src tyrosine kinase associates with the catalytic domain of ErbB-2: implications for ErbB-2 mediated signaling and transformation. Oncogene 24, 7599-7607.

Kobayashi, S., Boggon, T. J., Dayaram, T., Jänne, P. A., Kocher, O., Meyerson, M., Johnson, B. E., Eck, M. J., Tenen, D. G., and Halmos, B. (2005). EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. New England Journal of Medicine *352*, 786-792.

Kubbutat, M. H., Jones, S. N., and Vousden, K. H. (1997). Regulation of p53 stability by Mdm2. Nature *387*, 299-303.

Lai, E. C. (2004). Notch signaling: Control of cell communication and cell fate. Development *131*, 965-973.

Lakin, N. D., and Jackson, S. P. (1999). Regulation of p53 in response to DNA damage. Oncogene *18*, 7644-7655.

Lemmon, M. A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. Cell *141*, 1117-1134.

Leong, K. G., and Gao, W. Q. (2008). The Notch pathway in prostate development and cancer. Differentiation *76*, 699-716.

Leong, K. G., and Karsan, A. (2006). Recent insights into the role of Notch signaling in tumorigenesis. Blood *107*, 2223-2233.

Levene, P. A., and Schormueller, A. (1933). The synthesis of tyrosinephosphoric acid. J Biol Chem *100*, 583-587.

Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. Cell *88*, 323-331.

Li, S. C., Sonoyang, Z., Vincent, S. J. F., Zwahlen, C., Wiley, S., Cantley, L., Kay, L. E., Forman-Kay, J., and Pawson, T. (1997). High-affinity binding of the Drosophila Numb phosphotyrosine-binding domain to peptides containing a Gly-Pro-(p)Tyr motif. Proceedings of the National Academy of Sciences of the United States of America *94*, 7204-7209.

Li, S. C., Zwahlen, C., Vincent, S. J., McGlade, C. J., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (1998). Structure of a Numb PTB domain-peptide complex suggests a basis for diverse binding specificity. Nature structural biology *5*, 1075-1083.

Liu, B. A., Engelmann, B. W., and Nash, P. D. (2012). The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. FEBS Lett *586*, 2597-2605.

Liu, B. A., Jablonowski, K., Raina, M., Arce, M., Pawson, T., and Nash, P. D. (2006). The human and mouse complement of SH2 domain proteins-establishing the boundaries of phosphotyrosine signaling. Molecular cell *22*, 851-868.

Liu, B. A., Shah, E., Jablonowski, K., Stergachis, A., Engelmann, B., and Nash, P. D. (2011). The SH2 domain-containing proteins in 21 species establish the provenance and scope of phosphotyrosine signaling in eukaryotes. Sci Signal *4*, ra83.

Lu, Y., Zi, X., Zhao, Y., Mascarenhas, D., and Pollak, M. (2001). Insulin-Like Growth Factor-I Receptor Signaling and Resistance to Trastuzumab (Herceptin). JNCI Journal of the National Cancer Institute *93*, 1852-1857.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science *298*, 1912-1934.

McGill, M. A., Dho, S. E., Weinmaster, G., and McGlade, C. J. (2009). Numb regulates post-endocytic trafficking and degradation of notch1. Journal of Biological Chemistry *284*, 26427-26438.

McGill, M. A., and McGlade, C. J. (2003). Mammalian numb proteins promote Notch1 receptor ubiquitination and degradation of the Notch1 intracellular domain. J Biol Chem *278*, 23196-23203.

McPherson, P. S., Kay, B. K., and Hussain, N. K. (2001). Signaling on the endocytic pathway. Traffic *2*, 375-384.

Miaczynska, M., Pelkmans, L., and Zerial, M. (2004). Not just a sink: Endosomes in control of signal transduction. Current Opinion in Cell Biology *16*, 400-406.

Misquitta-Ali, C. M., Cheng, E., O'Hanlon, D., Liu, N., McGlade, C. J., Tsao, M. S., and Blencowe, B. J. (2011). Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. Mol Cell Biol *31*, 138-150.

Moll, U. M., and Petrenko, O. (2003). The MDM2-p53 interaction. Mol Cancer Res *1*, 1001-1008.

Momand, J., Zambetti, G. P., Olson, D. C., George, D., and Levine, A. J. (1992). The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. Cell *69*, 1237-1245.

Morooka, T., and Nishida, E. (1998). Requirement of p38 mitogen-activated protein kinase for neuronal differentiation in PC12 cells. J Biol Chem *273*, 24285-24288.

Morris, S. W., Kirstein, M. N., Valentine, M. B., Dittmer, K. G., Shapiro, D. N., Saltman, D. L., and Look, A. T. (1994). Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science *263*, 1281-1284.

Morrison, S. J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. Nature *441*, 1068-1074.

Mossé, Y. P., Laudenslager, M., Longo, L., Cole, K. A., Wood, A., Attiyeh, E. F., Laquaglia, M. J., Sennett, R., Lynch, J. E., Perri, P., *et al.* (2008). Identification of ALK as a major familial neuroblastoma predisposition gene. Nature *455*, 930-935.

Murray, P. B., Lax, I., Reshetnyak, A., Ligon, G. F., Lillquist, J. S., Natoli, E. J., Jr., Shi, X., Folta-Stogniew, E., Gunel, M., Alvarado, D., and Schlessinger, J. (2015). Heparin is an activating ligand of the orphan receptor tyrosine kinase ALK. Sci Signal *8*, ra6.

Nam, Y., Weng, A. P., Aster, J. C., and Blacklow, S. C. (2003). Structural requirements for assembly of the CSL·intracellular Notch1·Mastermind-like 1 transcriptional activation complex. Journal of Biological Chemistry *278*, 21232-21239.

Nie, J., Li, S. S. C., and McGlade, C. J. (2004). A novel PTB-PDZ domain interaction mediates isoform-specific ubiquitylation of mammalian numb. Journal of Biological Chemistry *279*, 20807-20815.

Nie, J., McGill, M. A., Dermer, M., Dho, S. E., Wolting, C. D., and McGlade, C. J. (2002). LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation. EMBO Journal *21*, 93-102.

Nishimura, T., Yamaguchi, T., Tokunaga, A., Hara, A., Hamaguchi, T., Kato, K., Iwamatsu, A., Okano, H., and Kaibuchi, K. (2006). Role of numb in dendritic spine development with a Cdc42 GEF intersectin and EphB2. Molecular biology of the cell *17*, 1273-1285.

Palmer, R. H., Vernersson, E., Grabbe, C., and Hallberg, B. (2009). Anaplastic lymphoma kinase: Signalling in development and disease. Biochemical Journal *420*, 345-361.

Pancewicz, J., and Nicot, C. (2011). Current views on the role of Notch signaling and the pathogenesis of human leukemia. BMC Cancer *11*, 502.

Pece, S., Confalonieri, S., P, R. R., and Di Fiore, P. P. (2011). NUMB-ing down cancer by more than just a NOTCH. Biochim Biophys Acta *1815*, 26-43.

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

Reynolds, A. B., Kanner, S. B., Bouton, A. H., Schaller, M. D., Weed, S. A., Flynn, D. C., and Parsons, J. T. (2014). SRChing for the substrates of Src. Oncogene *33*, 4537-4547.

Rhyu, M. S., Jan, L. Y., and Jan, Y. N. (1994). Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. Cell *76*, 477-491.

Roy, M., Pear, W. S., and Aster, J. C. (2007). The multifaceted role of Notch in cancer. Current Opinion in Genetics and Development *17*, 52-59.

Salcini, A. E., Confalonieri, S., Doria, M., Santolini, E., Tassi, E., Minenkova, O., Cesareni, G., Pelicci, P. G., and Di Fiore, P. P. (1997). Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. Genes Dev *11*, 2239-2249.

Santolini, E., Puri, C., Salcini, A. E., Gagliani, M. C., Pelicci, P. G., Tacchetti, C., and Di Fiore, P. P. (2000). Numb is an endocytic protein. The Journal of cell biology *151*, 1345-1352.

Sasaki, T., Rodig, S. J., Chirieac, L. R., and Janne, P. A. (2010). The biology and treatment of EML4-ALK non-small cell lung cancer. Eur J Cancer *46*, 1773-1780.

Schlessinger, J., Plotnikov, A. N., Ibrahimi, O. A., Eliseenkova, A. V., Yeh, B. K., Yayon, A., Linhardt, R. J., and Mohammadi, M. (2000). Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. Molecular cell *6*, 743-750.

Sefton, B. M., Hunter, T., Beemon, K., and Eckhart, W. (1980). Evidence that the phosphorylation of tyrosine is essential for cellular transformation by Rous sarcoma virus. Cell *20*, 807-816.

Shiota, M., Fujimoto, J., Semba, T., Satoh, H., Yamamoto, T., and Mori, S. (1994). Hyperphosphorylation of a novel 80 kDa protein-tyrosine kinase similar to Ltk in a human Ki-1 lymphoma cell line, AMS3. Oncogene *9*, 1567-1574.

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., *et al.* (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561-566.

Sousa, L. P., Lax, I., Shen, H., Ferguson, S. M., De Camilli, P., and Schlessinger, J. (2012). Suppression of EGFR endocytosis by dynamin depletion reveals that EGFR signaling occurs primarily at the plasma membrane. Proceedings of the National Academy of Sciences of the United States of America *109*, 4419-4424.

Stoica, G. E., Kuo, A., Aigner, A., Sunitha, I., Souttou, B., Malerczyk, C., Caughey, D. J., Wen, D., Karavanov, A., Riegel, A. T., and Wellstein, A. (2001). Identification of Anaplastic Lymphoma Kinase as a Receptor for the Growth Factor Pleiotrophin. Journal of Biological Chemistry *276*, 16772-16779.

Stoica, G. E., Kuo, A., Powers, C., Bowden, E. T., Sale, E. B., Riegel, A. T., and Wellstein, A. (2002). Midkine binds to anaplastic lymphoma kinase (ALK) and acts as a growth factor for different cell types. Journal of Biological Chemistry *277*, 35990-35998.

Tan, C. S. H., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M. O., Claus, J., Bader, G. D., Aebersold, R., Pawson, T., and Linding, R. (2009). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. Science Signaling *2*.

Toriya, M., Tokunaga, A., Sawamoto, K., Nakao, K., and Okano, H. (2006). Distinct functions of human numb isoforms revealed by misexpression in the neural stem cell lineage in the Drosophila larval brain. Developmental Neuroscience *28*, 142-155.

Uemura, T., Shepherd, S., Ackerman, L., Jan, L. Y., and Jan, Y. N. (1989). numb, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos. Cell *58*, 349-360.

Ullrich, A., and Schlessinger, J. (1990). Signal transduction by receptors with tyrosine kinase activity. Cell *61*, 203-212.

Ushiro, H., and Cohen, S. (1980). Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes. Journal of Biological Chemistry *255*, 8363-8365.

Verdi, J. M., Bashirullah, A., Goldhawk, D. E., Kubu, C. J., Jamali, M., Meakin, S. O., and Lipshitz, H. D. (1999). Distinct human NUMB isoforms regulate differentiation vs.

proliferation in the neuronal lineage. Proceedings of the National Academy of Sciences of the United States of America *96*, 10472-10476.

Vieira, A. V., Lamaze, C., and Schmid, S. L. (1996). Control of EGF receptor signaling by clathrin-mediated endocytosis. Science *274*, 2086-2089.

Visser Smit, G. D., Place, T. L., Cole, S. L., Clausen, K. A., Vemuganti, S., Zhang, G., Koland, J. G., and Lill, N. L. (2009). Cbl controls EGFR fate by regulating early endosome fusion. Sci Signal *2*, ra86.

Volinsky, N., and Kholodenko, B. N. (2013). Complexity of receptor tyrosine kinase signal processing. Cold Spring Harb Perspect Biol *5*, a009043.

Westhoff, B., Colaluca, I. N., D'Ario, G., Donzelli, M., Tosoni, D., Volorio, S., Pelosi, G., Spaggiari, L., Mazzarol, G., Viale, G., *et al.* (2009). Alterations of the Notch pathway in lung cancer. Proc Natl Acad Sci U S A *106*, 22293-22298.

Wiesner, T., Lee, W., Obenauf, A. C., Ran, L., Murali, R., Zhang, Q. F., Wong, E. W., Hu, W., Scott, S. N., Shah, R. H., *et al.* (2015). Alternative transcription initiation leads to expression of a novel ALK isoform in cancer. Nature *526*, 453-457.

Wiley, H. S., Herbst, J. J., Walsh, B. J., Lauffenburger, D. A., Rosenfeld, M. G., and Gill, G. N. (1991). The role of tyrosine kinase activity in endocytosis, compartmentation, and down-regulation of the epidermal growth factor receptor. Journal of Biological Chemistry *266*, 11083-11094.

Wilkin, M. B., and Baron, M. (2005). Endocytic regulation of Notch activation and down-regulation (review). Molecular Membrane Biology *22*, 279-289.

Yamamoto, S., Charng, W. L., and Bellen, H. J. (2010). Endocytosis and intracellular trafficking of notch and its ligands. In Current Topics in Developmental Biology, pp. 165-200.

Yarden, Y., and Sliwkowski, M. X. (2001). Untangling the ErbB signalling network. Nat Rev Mol Cell Biol *2*, 127-137.

Yoshida, T., Tokunaga, A., Nakao, K., and Okano, H. (2003). Distinct expression patterns of splicing isoforms of mNumb in the endocrine lineage of developing pancreas. Differentiation *71*, 486-495.

Zhang, K., Duan, L., Ong, Q., Lin, Z., Varman, P. M., Sung, K., and Cui, B. (2014). Light-mediated kinetic control reveals the temporal effect of the Raf/MEK/ERK pathway in PC12 cell neurite outgrowth. PLoS One *9*, e92917.

Zhong, W., Feder, J. N., Jiang, M. M., Jan, L. Y., and Jan, Y. N. (1996). Asymmetric localization of a mammalian numb homolog during mouse cortical neurogenesis. Neuron *17*, 43-53.

Zhong, W., Jiang, M. M., Weinmaster, G., Jan, L. Y., and Jan, Y. N. (1997). Differential expression of mammalian Numb, Numblike and Notch1 suggests distinct roles during mouse cortical neurogenesis. Development *124*, 1887-1897.

Zhou, B. P., Liao, Y., Xia, W., Zou, Y., Spohn, B., and Hung, M. C. (2001). HER-2/neu induces p53 ubiquitination via Akt-mediated MDM2 phosphorylation. Nat Cell Biol *3*, 973-982.

Zhou, P., Alfaro, J., Chang, E. H., Zhao, X., Porcionatto, M., and Segal, R. A. (2011). Numb links extracellular cues to intracellular polarity machinery to promote chemotaxis. Dev Cell *20*, 610-622.

Zwahlen, C., Li, S. C., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (2000). Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. The EMBO journal *19*, 1505-1515.

Chapter 2

2 Investigation of NUMB-PTB interactome

2.1 Abstract

Cellular events rely on protein-protein interactions that are often mediated by the modular domains which recognize particular sequence motifs within their binding partners. The NUMB protein is the first described cell fate determinant that is conserved from fly to human. Further studies have since revealed the complex nature of NUMB functions that are involved in a wide range of cellular events. As a tumor suppressor, NUMB attenuates the activity of NOTCH which is one of the most deregulated receptors in cancer, and stabilizes p53 from ubiquitination and degradation. NUMB mainly mediates protein interactions via its modular PTB domain. Here we present a systematic investigation of the NUMB-PTB interactome by employing an integrative strategy combining both protein array and peptide array. NUMB-PTB binding specificity is thoroughly profiled and explained at a structural level. Interestingly, we find that RTKs are highly enriched in the NUMB interactome. The interactions with select RTKs are further verified *in vivo*, including ALK, ErbB2 and FGR. Our study not only provides a systematic overview of the NUMB-PTB interactome, but also reveals novel interactions between NUMB and RTKs that have potential clinical value for the diagnosis and treatment of RTK positive cancers.

2.2 Introduction

Proteins involved in many important cellular events are often found to be interacting with each other. Protein-protein interactions (PPIs) play a pivotal role in biological processes and mediate signaling transduction that is often mis-regulated in many cancers (Arkin and Wells, 2004). Characterizing the interactome of a certain protein can be crucial to systematically understand its functions. NUMB is the first defined cell fate determinant, originally identified in *Drosophila* (Rhyu et al., 1994; Uemura et al., 1989). NUMB plays an important role in the development of both *Drosophila* and mammals, especially in the development of the neural system. NUMB-deficient mice die at an early embryonic stage and display severe neural defects (Zhong et al., 2000). At the molecular level, NUMB exhibits a complex array of functions and is involved in a multitude of biological processes including ubiquitination-mediated protein degradation, endocytosis, cell adhesion, cell polarity, cell migration and tumorigenesis (Gulino et al., 2010). This wide range of roles is likely due to the heterogeneity of NUMB-interacting proteins from different genetic backgrounds.

Most NUMB functions are uncovered during identification of novel NUMB-binding proteins. For instance, NUMB was found to play a role in the ubiquitin network by directly interacting with several E3 ligases: Mdm2 (Juven-Gershon et al., 1998), Lnx (Nie et al., 2002), Itch (Di Marcotullio et al., 2006) and Siah1 (Susini et al., 2001). NUMB has also been defined as an endocytic protein because it interacts with endocytic regulators EH, EPS15, EPS15R and AP-2 (Berdnik et al., 2002; Salcini et al., 1997) and localizes to the cell membrane peripheral region (Dho et al., 1999). During endocytosis, NUMB mediates the internalization of a number of membrane associated proteins, including NOTCH, Integrins, E-cadherin and TrkB, all of which were discovered through their interactions with NUMB (McGill et al., 2009a; Nishimura and Kaibuchi, 2007; Sato et al., 2011; Wang et al., 2009; Zhou et al., 2011). Similarly, the function of NUMB in cell polarity and migration was discovered alongside the identification of NUMB-Par complex (Par3/Par6/aPKC) interaction (Nishimura and Kaibuchi, 2007; Sato et al., 2011; Wang et al., 2009).

As an adaptor protein, NUMB contains a prototypical protein-protein interaction domain, the phosphotyrosine binding (PTB) domain, which is evolutionarily and functionally conserved. Based on the SMART database, there exists at least 5208 PTB domains within 4530 proteins across different species. In particular, the human genome encodes 46 proteins containing 51 PTB domains. PTB domain-containing proteins are often adaptor and scaffold proteins, which organize and regulate the signaling networks involved in a wide variety of biological processes (Uhlik et al., 2005). PTB domains can be classified into three groups: IRS-1/Dok-like, Shc-like, and Dab-like (Uhlik et al., 2005). The IRS-1/Dok-like and Shc-like groups are phosphotyrosine-dependent; they specifically bind to p-Tyr sites similar to SH2 domain containing proteins. The Dab-like group also recognizes tyrosine residues, but the binding is phosphotyrosine independent. NUMB-PTB is one of many PTB domains that belong to the Dab-like group. Even though they are not directly involved in tyrosine phosphorylation mediated signaling transduction, Dab-like PTB containing proteins participate in the regulation of various cellular events such as endocytosis (Howell et al., 1999; Kinoshita et al., 2001), cell membrane protein processing (Guenette et al., 1999; Hill et al., 2003), asymmetric cell division (Chien et al., 1998) and integrin-mediated cell adhesion (Chang et al., 1997; Filardo et al., 1995).

Considering the complexity of its functional pattern, it is not surprising that NUMB has been implicated in cancer as a tumor suppressor. The cancer-related role of NUMB mainly relies on the interplay between NUMB and either the NOTCH receptor, or MDM2/p53 complex. NUMB directly binds to NOTCH (Guo et al., 1996) and p53 (Dhami et al., 2013) through its PTB domain but regulates NOTCH and p53 via different mechanisms. NUMB promotes the internalization and subsequent degradation of cell membrane-localized NOTCH, thus attenuating NOTCH-stimulated signaling of cell proliferation (McGill et al., 2009b; McGill and McGlade, 2003). In contrast, NUMB stabilizes p53 activity by preventing MDM2-mediated ubiquitination and degradation of p53 (Colaluca et al., 2008). Both NUMB-NOTCH and NUMB-p53 interactions bear great potential for the development of novel anti-cancer targeted therapeutics, whose functions and applications can be elucidated through a systematic investigation of the NUMB interactome. Conventionally, a number of high-throughput methods have been developed for mapping the interactome, such as affinity-purification mass spectrometry (AP/MS), and the hybrid-

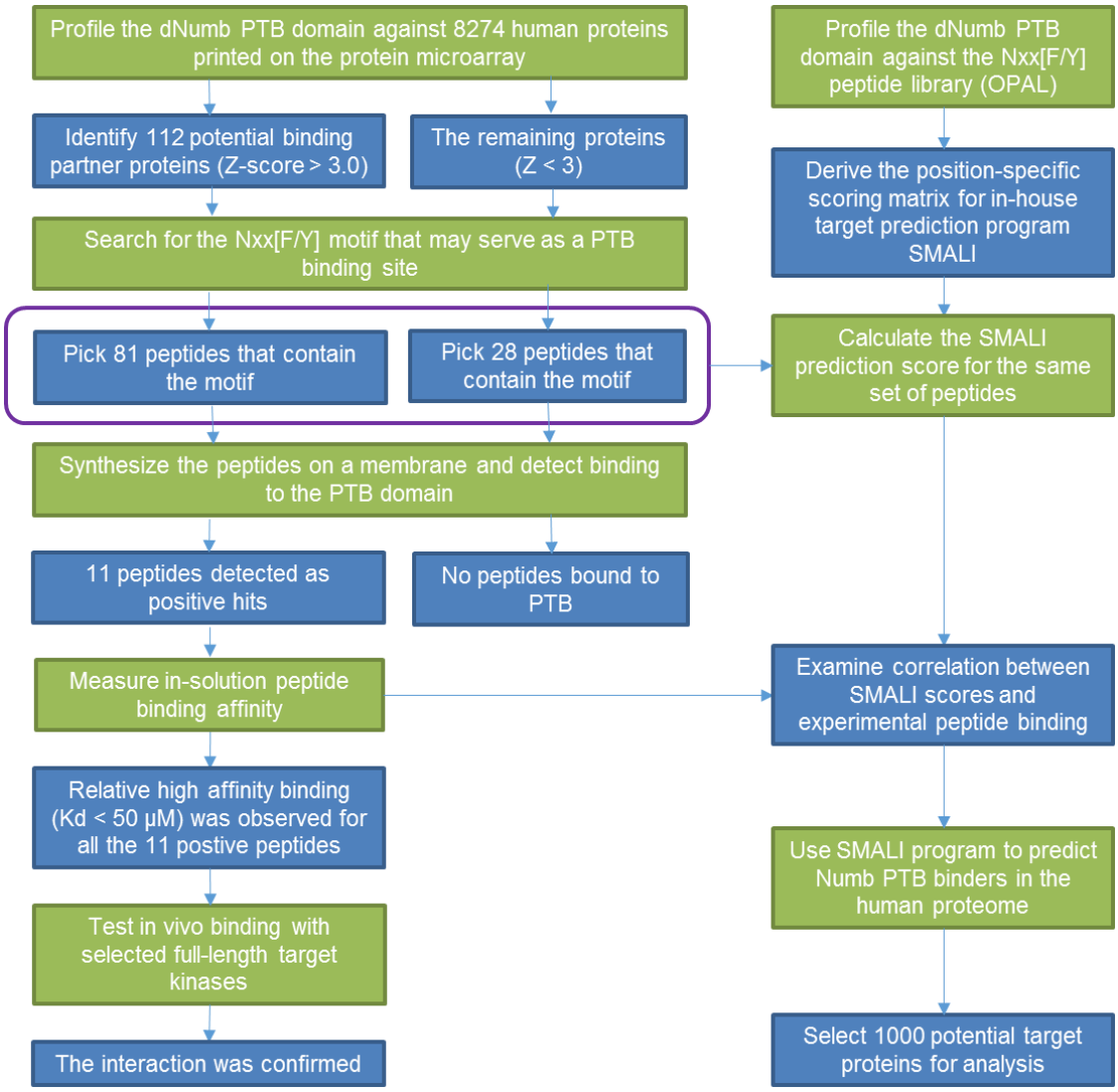


Figure.2.1 A strategy to identify NUMB-PTB binding specificity and interactome

The Green text boxes represent operation, while the blue text boxes represent output. A commercially available ProtoArray was probed with purified NUMB PTB domain. In complementary experiments, the specificity profile of NUMB PTB, as determined by OPAL peptide array, was used to score the NxxF/Y motifs in the pool of proteins identified by ProtoArray using the computer program SMALI. NxxF/Y motifs scored above a cut off value were further validated *in vitro* and *in vivo*.

reporting systems: yeast two hybrid (Y2H). Although AP/MS is one of the most powerful approaches for PPI studies, the technique does not distinguish between direct and indirect interactions or detect low-abundance proteins. Y2H is useful *in vivo* system, but is less time- and cost-efficient. Here we represent an alternative strategy for effectively mapping the interactome of the NUMB PTB domain (Fig.2.1). Briefly the direct NUMB-PTB

interactors were profiled by probing biotinylated NUMB-PTB on a commercially available protein array, which contains 8274 purified and normalized human proteins. Simultaneously, an oriented peptide array library (OPAL) (Rodriguez et al., 2004) was synthesized and probed with the NUMB-PTB. The data was processed using the SMALI (scoring matrix-assisted ligand identification) program (Li et al., 2008a) to generate a prediction of NUMB-PTB binding specificity and sites. The PTB-binding proteins and sites were further validated using protein-peptide binding assay, immunoprecipitation and pull down assay. Intriguingly, we identified several receptor tyrosine kinases (RTK) as PTB direct binding partners both *in vitro* and *in vivo*, which had never been reported before. In addition, when the PTB binding site predictions were applied across the whole human genome, a number of RTKs were high-scored and ranked at the top of the list. One such interaction occurs between NUMB and ALK, a receptor tyrosine kinase involved in tumorigenesis (presented in Chapter 3). NUMB mediates ALK internalization and regulates its post-endosomal trafficking antagonistically in an isoform-dependent manner.

2.3 Material and methods

2.3.1 Cell culture

HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 µg/ml penicillin, 100 µg/ml streptomycin and 2 mg/ml L-glutamine. SKBR3 cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% FBS, 100 µg/ml penicillin, 100 µg/ml streptomycin and 2 mg/ml L-glutamine. HEK293/ALK stable cells were grown in DMEM supplemented with 10% FBS, 100 µg/ml penicillin, 100 µg/ml streptomycin, 2 µg/ml puromycin and 2 mg/ml L-glutamine. Cells were incubated at 37 °C in a humidified atmosphere containing 5% carbon dioxide.

2.3.2 GST-NUMB-PTB expression, purification and biotinylation

The BL21 strain of *Escherichia coli* was transformed with pGEX6P3-dNUMB-PTB or pGEX6P3-mNUMB-PTB^{L}/PRB^{S}. Positive colonies were cultured in Lysogeny Broth (LB) medium to a density of OD_{600} 0.6~0.8, then the protein expression was induced with 0.5 mM IPTG for 16 hours at 18°C. The bacterial cells were harvested and pellets were re-suspended in PBS buffer containing complete protease inhibitors (Roche). Triton X-100 was added to a final concentration of 2%, lysozyme was added to a final concentration 1 mg/ml and benzonase was added to a final concentration 20 units/ml. The suspension was sonicated six times (10 seconds each) on ice and then lysed for 30 min at room temperature. Lysates were centrifuged at 15,000 × g for 30 min at 4°C and the supernatant was collected. Purification of GST-tagged proteins was performed with glutathione resin (GE healthcare). The resin was washed with PBS buffer three times. The lysate supernatant was loaded to the resin followed by three column volume washes with PBS buffer, then eluted with 10 mM glutathione in PBS buffer (no elution step for pulldown assay). To determine GST-protein purity, a small amount of purified protein (or protein on resin) was boiled with SDS-loading dye and then analyzed by SDS-PAGE/Coomassie staining. The GST NUMB-PTB domains used for the Protoarray hybridization were biotin-labeled using EZ-Link™ Sulfo-NHS-Biotinylation Kit (Thermo) in 1×PBS buffer.

2.3.3 ProtoArray hybridization

The Protoarray Human microarray (Invitrogen) was equilibrated at room temperature for 30 minutes prior to blocking. The array was arranged barcode-side up in each well of a chilled 4-chamber incubation tray. 5 ml blocking buffer (50 mM HEPES pH 7.5, 200 mM NaCl, 0.08% Triton X-100, 25% Glycerol, 20 mM Reduced glutathione, 1 mM DTT and 1× Synthetic Block) was then added to the array and incubated for 1 hour at 4 °C with slow shaking. After blocking buffer was removed, biotinylated GST-dNUMB-PTB was diluted to final concentration of 5 µM in washing buffer (1×PBS, pH 7.4, 1×Synthetic Block, 0.1% Tween 20) to form the blotting solution. 120 µL of blotting solution was added on top of the array surface and then a LifterSlip cover glass was carefully attached to the array. The array was blotted for 90 minutes at 4 °C without shaking. The cover glass was then removed and the array was washed with washing buffer 5 times for 5 minutes each at 4 °C with slow shaking. The array was visualized by incubation with 5 ml Alexa Fluor 647-conjugated streptavidin for 90 minutes at 4 °C with gentle shaking. Then the array was washed again following the same procedures as above. Finally, the array was dipped into distilled water once to remove salt residues and scanned using Tecan fluorescence microarray scanner. Data were analyzed using ArrayPro software.

2.3.4 Transient transfection

Plasmid DNA was first diluted in serum-free Minimum Essential Media (MEM) to a final concentration of 0.01 µg/µl. For each 10-cm dish, 40µl X-tremeGENE HP DNA transfection reagent (Roche) was mixed with 1ml plasmid DNA in MEM. The mixture was added to the culture dish (10ml) after 15 min incubation at room temperature. After transfection, cells were incubated for 48~72 hours to allow for ectopic gene expression.

2.3.5 Immunoprecipitation, pull down and western blotting

Cultured cells were lysed in ice-cold mammalian cell lysis buffer (1% Triton X-100, 50 mM Tris-pH 7.2, 150 mM NaCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM EGTA, 0.5 mM DTT, 10 mM NaF) containing complete protease inhibitors (Roche). Cell debris was removed by centrifugation, and 500 µg of supernatant protein was incubated for four hours

in the presence of 1 µg antibody and 30 µl of 50% slurry protein G-Sepharose beads (Roche), or 1 µg immobilized GST fusion protein (beads). The beads were subsequently washed three times with lysis buffer and boiled with SDS-loading dye. The precipitates were resolved by SDS-PAGE. Proteins were then transferred to polyvinylidene fluoride PVDF membrane using semi-dry transferring method, and detected by immunoblotting with appropriate antibodies and visualized by enhanced chemi-luminescence (ECL).

2.3.6 Free peptide and peptide array synthesis

Both free and membrane-bound peptides were synthesized using an automatic Intavis AG workstation from amino acid monomers protected with Fmoc (9-fluorenylmethyl-oxycarbonyl). Rink-resin (Rink-NH_2) was used to couple the first amino acid in free peptide synthesis, whereas the amine-derivate cellulose membrane (cellulose-NH_2) was made to couple the first amino acid in on-membrane peptide synthesis. In each synthesis cycle, the carboxyl group of Fmoc-protected amino acid (Fmoc-R-COOH) was first linked to the amine group of the previous amino acid (or Rink-NH_2) through an amide bond. All unoccupied amine groups were then blocked (acetylated) by acetic anhydride to prevent incorrect amide bonds forming in subsequent cycles. Next, the Fmoc group was removed by piperidine (de-protecting) to release the free amine group primed for linking the carboxyl group of next amino acid residue. Fluorescein-NHS was linked to the amine group of the last amino acid in free peptide synthesis. After synthesis, the on-membrane peptides were treated with a mixture containing 47.5% TFA (trifluoroacetic acid), 1.5% TIPS (tri-isopropylsilane) and 51% water, to remove all other protecting groups on amino acid side chains. The free peptides were treated with a mixture containing 95% TFA, 3% TIPS and 2% water to de-protect side chains while cleaving the peptide from Rink-resin.

2.3.7 Far western assay for peptide membrane blotting

The peptide membrane was blocked with 5% skim milk in TBST buffer (0.1 M Tris-HCl, pH 7.4, 150 mM NaCl, and 0.1% Tween 20) for 1 hour at room temperature, with slow shaking. GST-fusion proteins were added directly into the blocking buffer to reach a final concentration of 1µg/ml and incubated with the peptide array membrane for 1 hour at room

temperature, with slow shaking. The membrane was then washed with TBST buffer three times for 5 minutes each with shaking. Next, anti-GST-HRP antibody was added at 1:3000 (v/v) into 5% milk in TBST buffer. The membrane was incubated for 1 hour at room temperature with slow shaking followed by the same washing procedures as above. Finally, the bound GST-protein was visualized using ECL.

2.3.8 Determination of dissociation constant

The dissociation constant between protein and peptide was determined using a fluorescence polarization (FP) assay. Proteins and fluorescein labelled peptides were prepared following the procedures introduced above. To reduce variation, two independent tests were prepared simultaneously. 384-well flat bottom plates (Corning-3537) were used for the assay and a PerkinElmer Envision 2103 plate reader was used for evaluating FP. The peptides (approximate 2 µM each, mass) were dissolved in 100 µl DMSO and 3µl was diluted 100 fold in water. The concentrations of the proteins were determined and adjusted to 50 µM. 2-fold serial dilutions were set up to result in 16 different protein concentrations ranging from 0 to 50 µM. 30 µl of each protein sample was mixed with 5 µl peptide in the 384-well plate and the FP was measured. Non-linear fit was approximated by assuming peptide concentration was much lower than Kd value in the following equation:

$$\Delta FP = FP_{obs} - FP_0 = FP_{max} \times [\text{domain}] / (K_D + [\text{domain}])$$

The full equation without the approximation can be found in this study (Kaushansky et al., 2010).

2.3.9 Consensus binding motif and SMALI matrix

The OPAL membrane was scanned and the intensity was quantified on a Bio-Rad imaging system. To generate consensus motifs based on specific binding signals, the average background signal of the membrane was subtracted from the dataset. Since a conspicuous column of Lys or Arg was detected in the corresponding OPAL screen, these strong Lys or Arg signals are position-independent. To minimize position-independent effect, the

value for each Lys or Arg spot was readjusted by subtracting the average of the Lys or Arg in the column. See more details for SMALI scoring in (Li et al., 2008b).

2.4 Results

2.4.1 Screening NUMB PTB domain binding proteins using protein array

Similar amounts of protein are presented in each spot on the ProtoArray, creating equalized conditions for NUMB-PTB binding regardless of protein expression in different genetic backgrounds. The ProtoArray consists of 48 (12×4) grids, and each grid contains 400 (20×20) spots (Suppl.S1). The ProtoArray was incubated with 5 µM of biotinylated GST-NUMB-PTB domain followed by labelling with Alexa Fluor 647 streptavidin, and then the signal of hybridization was measured using a fluorescent micro-slide reader as shown in Fig.2.2A. Only a small portion of spots yielded a detectable signal, indicating the positive binders. The intact datasets for all grids were quantified and analyzed by ArrayProxy software.

The Z-scores were calculated for the signal intensity of each individual spot. To further examine the NUMB-PTB binding candidates, a minimum cut-off Z-score of 3.0 was applied (Suppl.S2), resulting in a list of 112 potential NUMB PTB binding partners (Suppl.S3). The candidate list was analyzed using String software (Szklarczyk et al., 2015) to map a functional protein association network. Next, the network was imported into Navigator software (Brown et al., 2009) to create a NUMB interacting network. As shown in Fig.2.2B, it is intriguing that the majority of candidates are protein kinases, including 30 tyrosine kinases and 34 serine/threonine kinases. Within the 64 kinases, PRKCQ, PRKCG, PPKCPE, RKCB, PRKCA and CAMK2B have been verified as NUMB direct interacting partners in previous reports (Martin-Blanco et al., 2014; Smith et al., 2007; Tokumitsu et al., 2005).

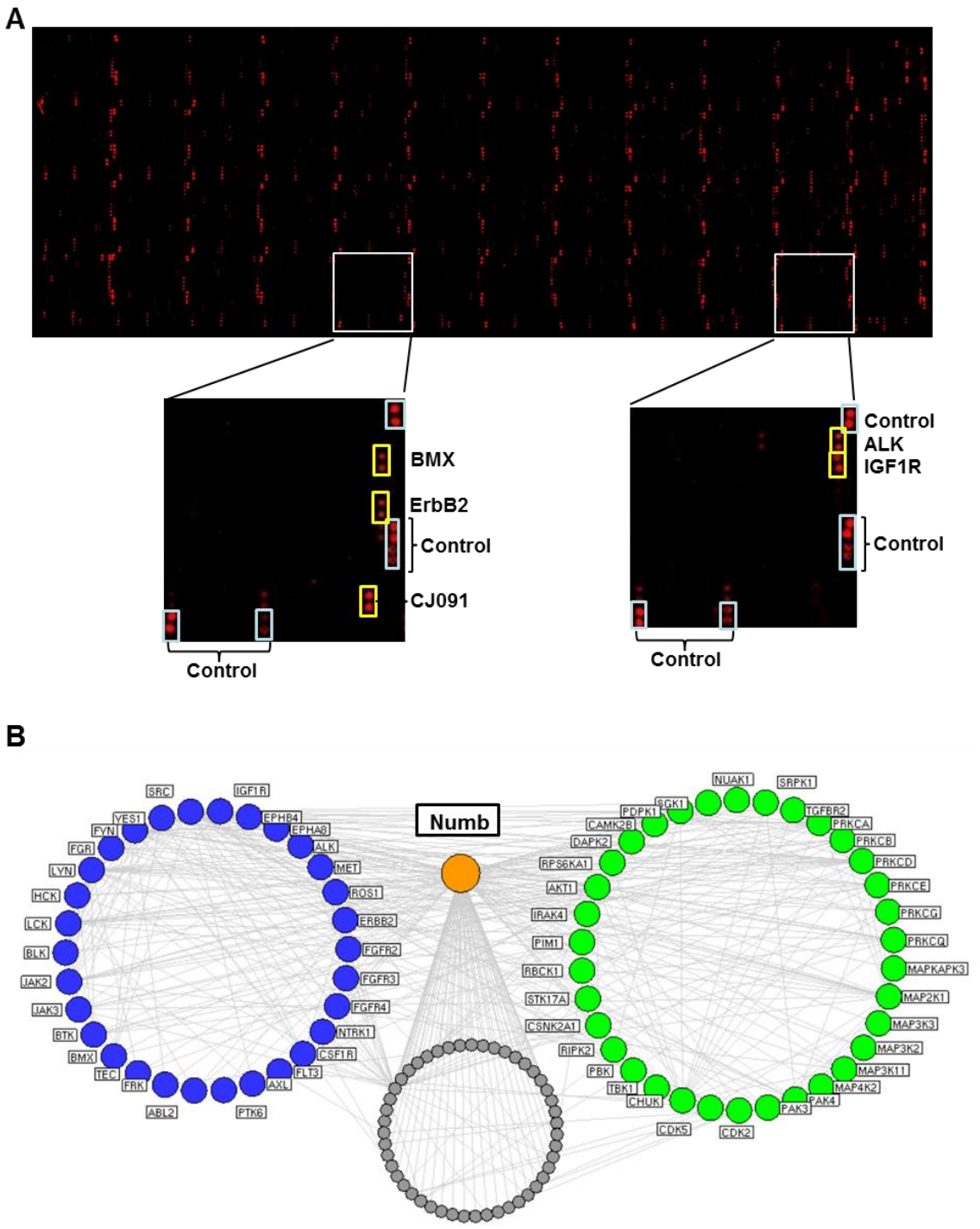


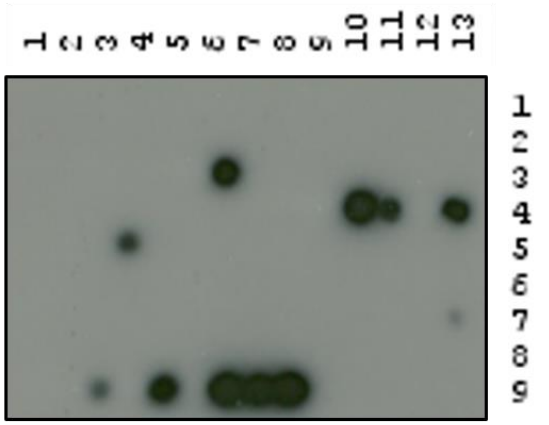
Figure.2.2 NUMB PTB domain interactome

(A) A protein array (ProtoArray®, Invitrogen) representing 8274 human proteins was probed with biotinylated dNUMB-PTB domain. (B) Proteins in the interactome of NUMB-PTB are grouped according to GO annotations. Color codes used are as follows: purple, tyrosine kinases; green, serine/threonine kinases; grey, all other NUMB-PTB binding partners.

2.4.2 Characterizing NUMB-PTB domain binding sites in the interactome

The NUMB PTB domain is known to bind NxxF/NxxY motifs (Li et al., 1998). To better evaluate our ProtoArray results, we selected 50 proteins which are disease especially cancer relevant from the highest Z-score 18.7 to as low as -0.3 and scanned their full sequences for NxxF and NxxY motifs. Peptides containing 11 amino acids for 109 motifs within the 50 protein we selected were synthesized on a nitrocellulose membrane, as well as 3 positive control peptides (9:6, JIP4_HUMAN; 9:7, NAK and 9:8, NAK-AA). The PTB domain of NUMB was expressed as a GST fusion protein in *E.coli* and purified. In a far western assay, the membrane was hybridized with GST-NUMB-PTB domain and anti-GST-HRP antibody sequentially, then illuminated by ECL solution and imaged using a BioRad imaging system (Fig.2.3A and Suppl.S4). 11 of 109 peptides exhibited detectable binding signals in this assay. Next, we synthesized free forms of the 11 peptides, as well as two positive control peptides IAIDNTAFMDE and GPIVNEGYVNT and two negative control peptides GKKGNLVYIID and LKPSNILYVDE. These peptides were attached with a fluorescein label and their binding affinities (dissociation constant, Kd) to NUMB-PTB were determined individually by using the fluorescence polarization (FP) assay. As shown in Fig.2.3B, the 11 positive peptides and positive control peptides have relatively high binding affinities to NUMB-PTB, whereas the negative control peptides exhibit no binding. All binding curves with Kd less than 50 μM are shown in Fig.2.4. It is interesting that peptides from three receptor tyrosine kinases, ALK, ErbB2 and ROS1, exhibited comparatively strong binding affinity to NUMB PTB domain, suggesting that NUMB may generally bind to receptor tyrosine kinases via the PTB domain. It should be noted that all proteins containing NUMB-PTB binding motifs, as verified in the far western assay, had Z-scores above 3.0 (Fig.2.6E and Suppl.S4), indicating that a reasonable cut-off was chosen in the data analysis of ProtoArray.

A



B

Protein Name	Uniprot Name	Peptide sequence	Kd(μM)	Position
NAK-AA	N/A	TGFSNAAFEDF	0.1±0.01	9:8
NAK	Q9U485_DROME	TGFSNMSFEDF	1.1±0.08	9:7
c10orf91	CJ091_HUMAN	QGFRNAAFEA	2.5±0.08	4:10
JIP4	JIP4_HUMAN	KGIENKAFDRN	0.9±0.05	9:6
LCK	LCK_HUMAN	RNLDNGGFYIS	40.2±4.3	9:4
ROS1	ROS1_HUMAN	SGVINESFEGE	0.7±0.02	3:6
ALK_1	ALK_HUMAN	GGHVNMAFSQS	0.7±0.04	4:13
ALK_2	ALK_HUMAN	FPCGNVNYGYQ	0.5±0.02	4:11
ALK_3	ALK_HUMAN	CGNVNYGYQQQ	1.6±0.15	5:3
FGR	FGR_HUMAN	SQAINPGFLDS	36.6±5	9:2
ERBB2	ERBB2_HUMAN	PAFDNLYYWDQ	9.4±0.3	7:13
Syntaxin-3	STX3_HUMAN	IAIDNTAFMDE	12±0.5	N/A
Synaptogrin-3	SNG3_HUMAN	GPIVNEGYVNT	13.7±0.8	N/A
CKIe	KC1E_HUMAN	GKKGNLVYIID	N/A	3:12
RSK-1	KS6A1_HUMAN	LKPSNILYVDE	N/A	8:6

Figure.2.3 Verifying NUMB-PTB domain binding sites in the interactome

(A) 108 NxxF and NxxY motifs, as predicted through the interactome, were synthesized in a peptide array and probed with NUMB-PTB domain (GST fusion). The horizontal numbers above the membrane represents the column number and the vertical numbers on the right of the membrane represents the row number; (C) The binding affinity of selected peptides identified in the peptide array, including motifs from tyrosine kinases ALK, ErbB2, FGR, LCK and ROS1.

Based on the results of far western and FP assays, we performed sequence logo analyses to study the amino acid preference of NUMB-PTB domain binding motifs. 13 peptides with binding affinity stronger than 50 μM were set as positive binding peptides of the NUMB-PTB domain in this analysis, while the remaining two peptides served as negative binding peptides. An initial one-sample analysis of all 112 peptides showed amino acid preferences only for N at -3 position and Y or F at zero position (Fig.2.5A). Next, we did

a positive and negative Two Sample Sequence Logo analysis, and determined a [G][F][V]Nx[A/G]Y/F[E][D] amino acid consensus sequence (Fig.2.5B). Fig.2.5B exhibits amino acid preferences in individually tested positive peptides. In Fig.2.5C, the amino acid residues are colored in a gradient from small (yellow) to large (green). There is an observed preference for small amino acids in the -6 and -1 positions. In Fig.2.5D, the amino acid residues are colored with red (acidic), blue (basic) and orange (aromatic) respectively, and it is apparent that acidic residues are preferred in the +1 and +2 positions while hydrophobic residues are preferred in the -5 position.

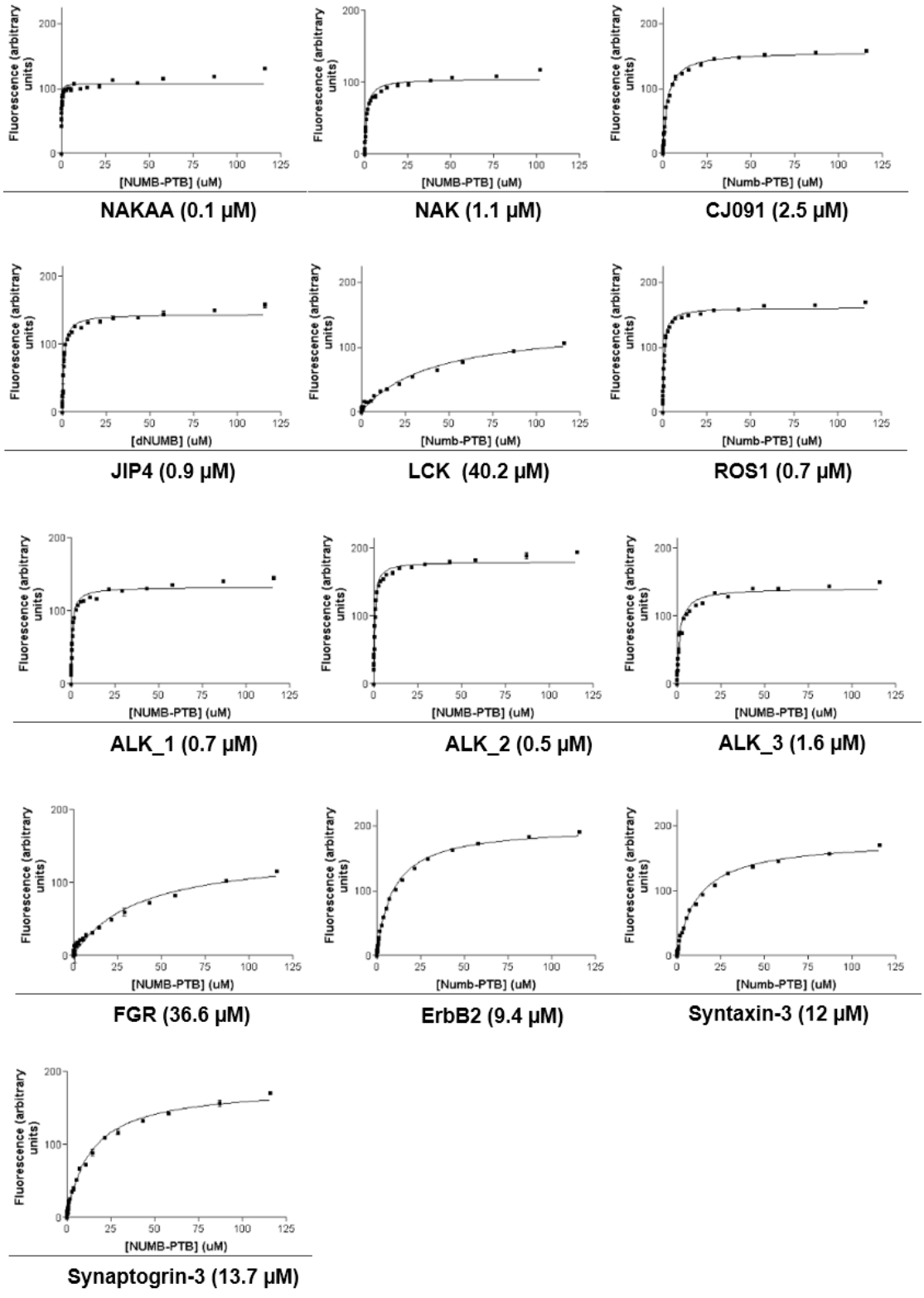


Figure.2.4 Binding curves for a selection of NUMB-PTB interactors

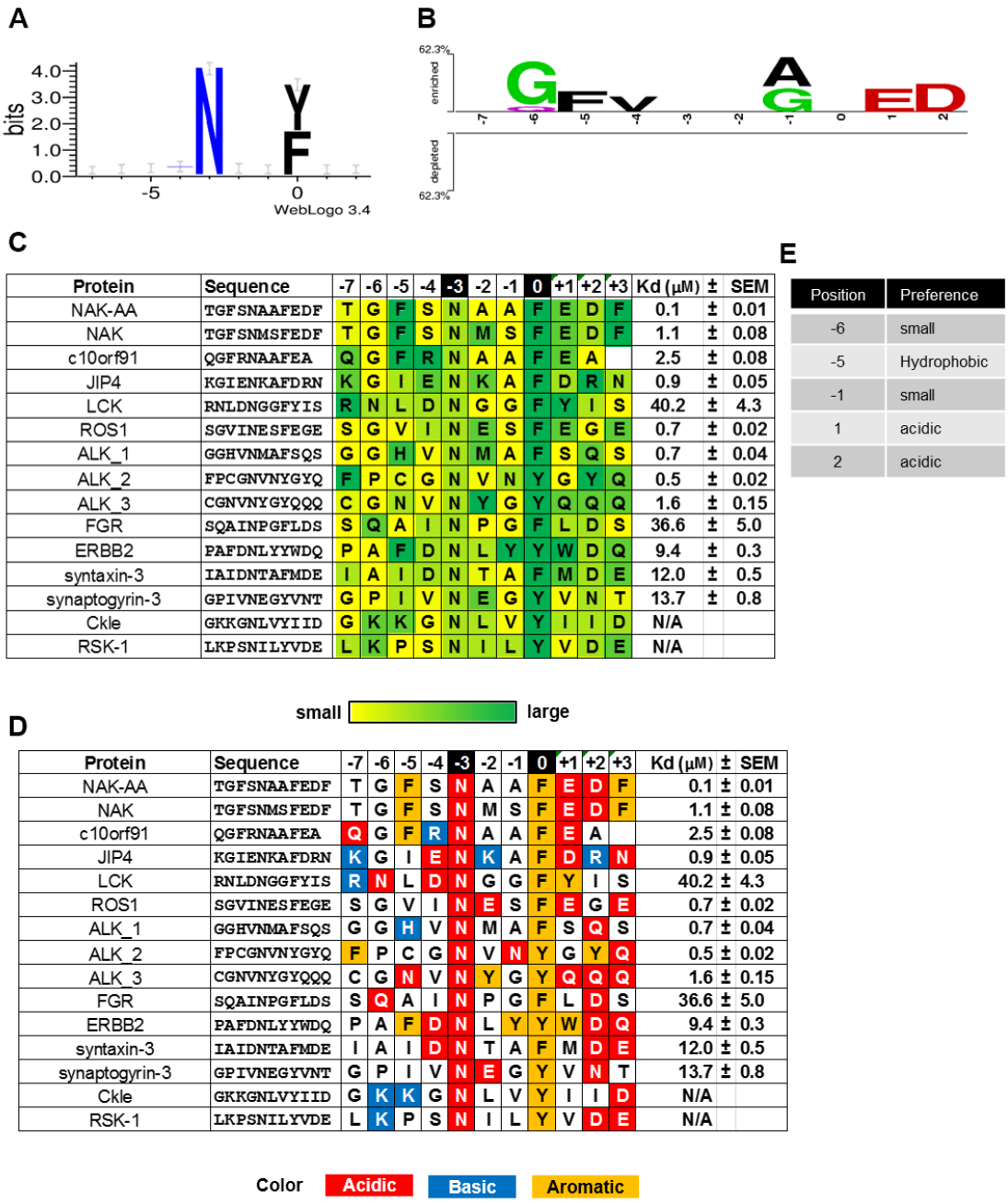


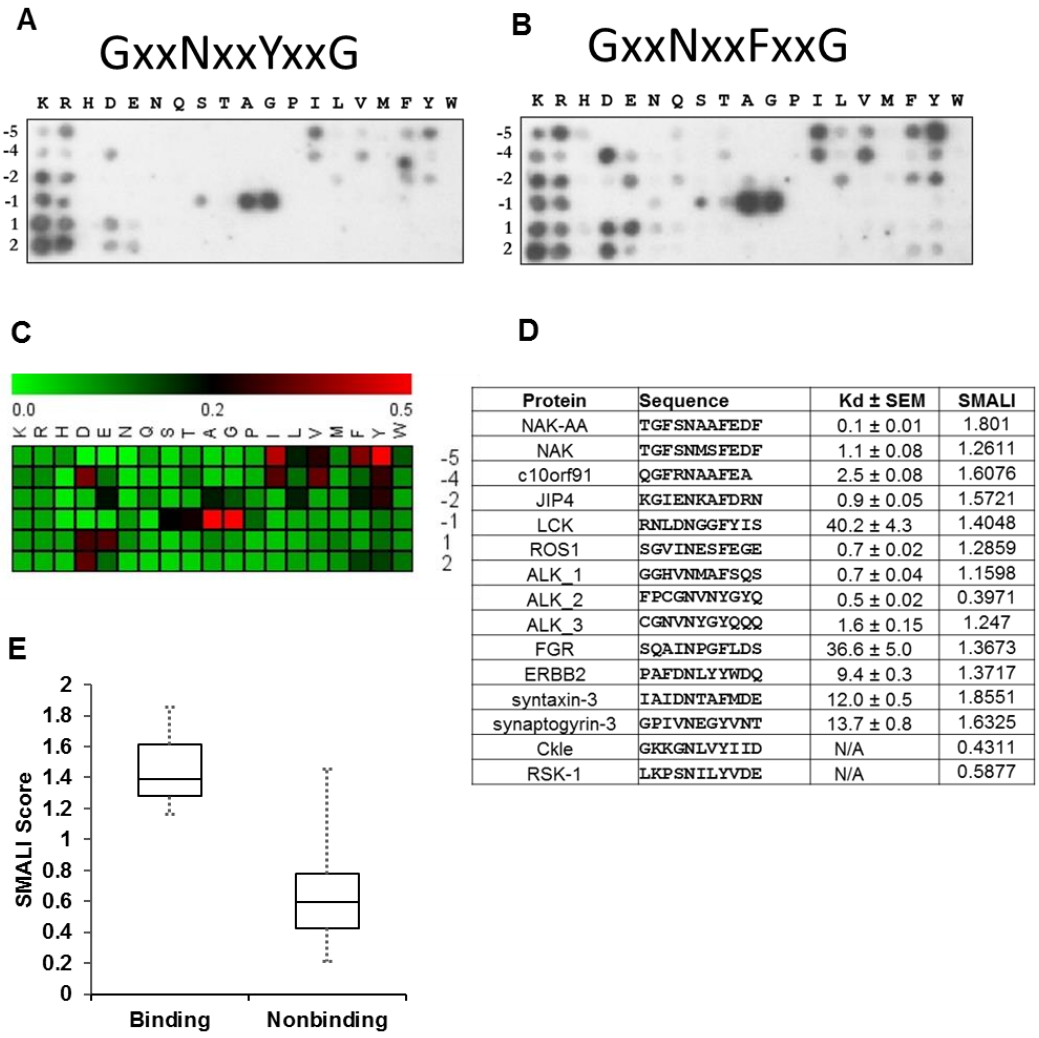
Figure.2.5 Amino acid preferences of verified NUMB-PTB binding motifs

(A-B) Sequence LOGOs were generated from motifs exhibiting high binding affinities to NUMB-PTB domain as determined by an *in vitro* FP assay. The residue positions are labeled from -7 to +2 with respect to F or Y, which is designated as position 0; (C) small amino acid residues at -6 and -1 positions are preferred for NUMB-PTB binding; (D) acid amino acidic residues are preferred at +1 and +2 positions while hydrophobic residues are preferred at -5 position; (E) overview of amino acid preferences in verified NUMB-PTB binding motifs.

2.4.3 Defining NUMB-PTB domain binding specificity

To systematically define the NUMB-PTB domain binding specificity, we employed an Oriented Peptide Array Library (OPAL) strategy. In parallel, we synthesized an OPAL array with N fixed at P-3 and P-0 positions directly on a nitrocellulose membrane. The flanking residues Gly (G) was added to each to mimic natural condition of the peptides. Recombinant GST-NUMB-PTB protein was incubated with this OPAL array and then detected using anti-GST-HRP. As shown in Fig.2.6A-B, the most obvious binding preferences based on the spot intensity value were observed at the P -1 position, where A and G were preferred. Weaker binding was observed at P-5 [I, Y, F], P-4 [D, I, V], P-2 [E, F, Y], P+1 [D, E] and P+2 [D] with even weaker signals at P-5 [V], P-4 [Y], P-2 [L], and P+2 [E]. The K/R spots were discarded, as these residues are known to display non-specific binding properties in peptide arrays. The array was quantified, by generating a scoring matrix that reflects the preference for an amino acid at a given position (Suppl.S5) and a heat map was constructed according to the scoring matrix (Fig.2.6C).

Next, we employed the SMALI approach to score the NUMB-PTB domain binding sites on the NUMB interactome identified above. A peptide with a higher SMALI score is considered to have a greater propensity for binding to NUMB-PTB domain. We first used SMALI to predict the binding between the 112 peptides and NUMB-PTB domain (Fig.2.6D and Suppl.S4). The SMALI scores for the positive (NUMB-PTB binding) peptides correlated well with their binding affinities (Kd values) and a significant difference was confirmed between the SMALI scores of the positive peptides and negative peptides (Fig.2.6E). 75% of positive peptides had SMALI scores greater than 1.25 while 75% of negative peptides have SMALI scores less than 0.75. Among the 11 positive peptides, the FPCGNVNYGYQ peptide has a relatively low SMALI score, which is due to the incorrect recognition of the NxxF motif in the SMALI program. This peptide has NVNYGY sequence containing partially overlapping NVNY and NVGY motifs. Indeed, the peptide shifted by two amino acids (CGNVNYGYQQQ) had a high SMALI score of 1.25, and we confirmed that NUMB-PTB bound to NYGY but not NVNY in this ALK peptide (Chapter 3). Overall, SMALI results represent an accurate prediction of NUMB-PTB binding sequences on the top hits of our ProtoArray data.



Protein	Sequence	Kd ± SEM	SMALI
NAK-AA	TGFSNAAFEDF	0.1 ± 0.01	1.801
NAK	TGFSNMSFEDF	1.1 ± 0.08	1.2611
c10orf91	QGFRNAAFEA	2.5 ± 0.08	1.6076
JIP4	KGIENKAFDRN	0.9 ± 0.05	1.5721
LCK	RNLDNGGFYIS	40.2 ± 4.3	1.4048
ROS1	SGVINESFEGE	0.7 ± 0.02	1.2859
ALK_1	GGHVNMAFSQS	0.7 ± 0.04	1.1598
ALK_2	FPCGNVNYGYQ	0.5 ± 0.02	0.3971
ALK_3	CGNVNYGYQQQ	1.6 ± 0.15	1.247
FGR	SQAINPGFLDS	36.6 ± 5.0	1.3673
ERBB2	PAFDNLYYWDQ	9.4 ± 0.3	1.3717
syntaxin-3	IAIDNTAFMDE	12.0 ± 0.5	1.8551
synaptogyrin-3	GPIVNEGYVNT	13.7 ± 0.8	1.6325
Ckle	GKKGNLVYIID	N/A	0.4311
RSK-1	LKPSNILYVDE	N/A	0.5877

Figure.2.6 Identifying NUMB-PTB binding specificity

(A-B) OPAL membranes GxxNxxYxxG and GxxNxxYxxG were probed with NUMB-PTB (GST fusion). Dark spots indicate positive binding; (C) Heat map visualization of amino acid preferences; (D) SMALI scores of verified binding and nonbinding motifs to the NUMB-PTB domain.

2.4.4 Structural explanation of NUMB-PTB binding specificity

Structure analysis was also performed to evaluate the binding preferences of the NUMB PTB domain. The structure of NUMB-PTB domain in complex with the NAK peptide (GFSNMSFEDFP) is shown in Fig.2.7A. (PDB code 1DDM, (Zwahlen et al., 2000)). The surface of NUMB PTB is colored according to its hydrophobicity while the peptide is shown in sticks. The PTB domain provides a hydrophobic pocket for binding peptides

between Gly at -6 position and Phe at position 0 (Fig.2.7A). The percentage of buried surface of each of the NAK peptide residues in NUMB PTB domain was calculated. A value of 100% means the ligand residue is fully buried in the interaction interface. Each residue at minus positions are captured well, with the exception of Met at -2 and Asp at +2 positions (Fig.2.7B). A close view of Ser at -1 position reveals that it is captured by a hydrophobic groove with no extra space to accommodate a larger amino acid. In addition, as pointed by Zwahlen et al (Zwahlen et al., 2000), a residue without chemical polarity would be more favorable by boosting the hydrophobic interaction, which supports the preference for A or G at this position (Fig.2.7C). Next, two PTB domain-peptide complex structures were used as parallel templates for modelling. The program MODELLER (Webb and Sali, 2014) was used to build a model structure.

The in-solution peptide binding assay indicated that Gly at -6 position is strongly preferred (Fig.2.5B). The NAK peptide in the dNUMB-PTB domain complex structure contains residues up to Gly at -6 position of the N-terminus but does not contain further residues. To study the potential role of Gly at -6 position of ligand peptides, we modelled an N-terminally extended NAK-AA peptide in complex with the dNUMB-PTB domain. The structure of the ARH (autosomal recessive hypercholesterolemia) PTB domain in complex with a 13-mer peptide (containing the residues from -8 to +4, PDB code 3SO6) was used as a secondary modelling template, in conjunction with the dNUMB-PTB domain complex structure. The sequence identity between the two PTB domains is 30% (37/123), and both ARH-PTB domain and Numb-PTB belong to the Dab-like PTB domain group (Fig.2.7D). The dNUMB PTB domain was then modelled in complex with an extended NAK peptide. Two additional magenta-colored residues are at the N-terminus of the peptide. The modelled peptide also contains two alanine replacements (Met to Ala at -2 position, Ser to Ala at -1 position), named NAK-AA, which has shown significantly increased binding affinity compared to the wild-type NAK peptide (Zwahlen et al., 2000). These two alanine residues increase the hydrophobic interaction (Fig.2.7E).

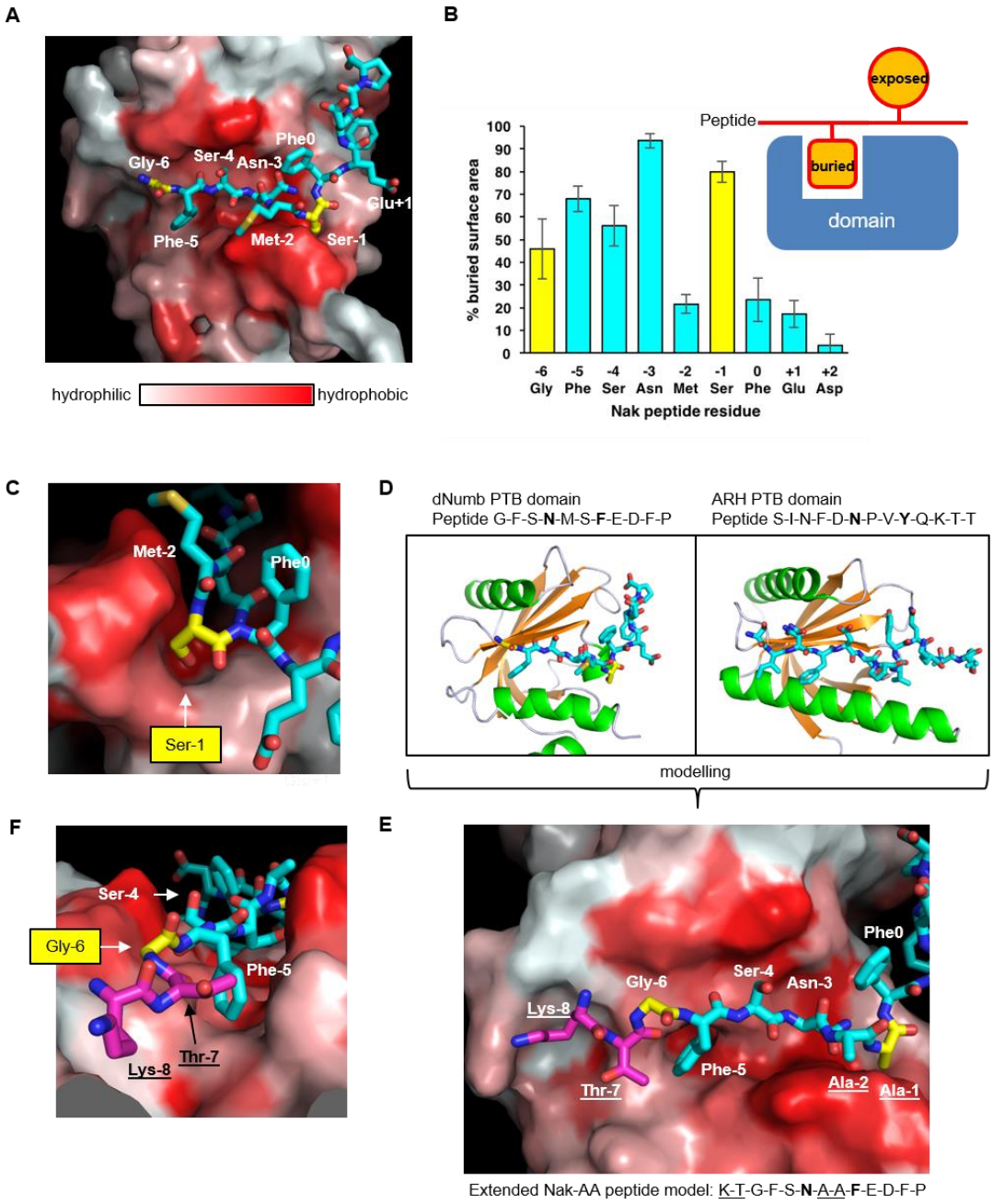


Figure.2.7 Structural analysis of NUMB-PTB binding specificity

(A) The NUMB-PTB domain provides a hydrophobic pocket for peptide binding between Gly at -6 position and Phe at position 0; (B) the percentage of buried surface for each residue in the NAK peptide in the complex; (C) a close-up-view of Ser at -1 position in the hydrophobic groove. This limited binding space likely explains why Ala or Gly is preferred at this position; (D) a comparison between NUMB-PTB binding and ARH PTB binding; (E) the preference of Ala in -1 and -2 positions contribute to the hydrophobic binding; (F) Gly at -6 position introduces a kink to help orient Thr at +7 and Lys at -8 positions in the groove without a steric hindrance.

In a close-up view of N-terminal residues in the modelled peptide, all residues sit well in the groove up to Lys at -8 position. The most likely role of the Gly at -6 position is to place Thr at -7 and Lys at -8 positions in the groove without steric hindrance, by introducing a kink at this position (Fig.2.7F).

2.4.5 Verification of NUMB RTK interactors *in vivo*

To confirm novel NUMB interactions *in vivo*, we performed pull down and/or co-immunoprecipitation assays. Three representative receptor tyrosine kinases, namely ALK, ErbB2, FGR whose deregulations are closely related to the abnormal behaviors of cancer cells were examined. We transiently expressed NUMB in HEK293 cells, which have inherently stable ALK expression. Reciprocal immunoprecipitations (IP) were performed between NUMB and ALK, and both proteins can be co-immunoprecipitated with each other (Fig.2.8A). The NUMB-ErbB2 interaction was verified by co-immunoprecipitation in SKBR3 cells which are known for ErbB2 overexpression and moderate NUMB expression (Fig.2.8B). The NUMB-FGR interaction was verified by pulldown assay with HeLa cell lysate (Fig.2.8C). We found FGR preferentially bound to the NUMB-PTB^{L} isoform, which is consistent with the fact that PTB^{L} isoforms are generally located in the cell membrane peripheral region where membrane-localized FGR also exists. Overall, this in vivo verification further raises the possibility that NUMB may act as a universal binding partner of RTKs.

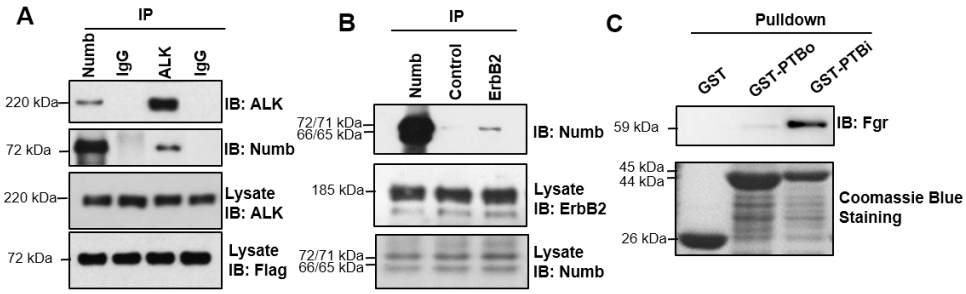


Figure.2.8 NUMB binds to ALK, ErbB2, and FGR receptors *in vivo*

(A) NUMB and ALK co-immunoprecipitate each other in HEK293 cells, which inherently express wildtype ALK. (B) ErbB2 co-immunoprecipitates NUMB in SKBR3 cells; (C) the long isoform of NUMB-PTB domain (PTB^{L}/PTBi) pull down more FGR compared to the short isoform (PTB^{S}/PTBo) in HeLa cells.

2.4.6 Predicting a genome-wide NUMB interactome by SMALI

The human genome consists of 20,000-25,000 genes, but expresses a much larger number of proteins once variants and isoforms are taken into account. In our study, the ProtoArray only represents a limited protein spectrum with 8274 proteins. Considering the excellent accuracy of our SMALI prediction for NUMB-PTB binding motifs, we scanned the entire human protein database (Uniprot) screening for more potential NUMB binding partners. The top 1000 peptide hits from 1083 proteins, which are classified according to GO annotations are shown in Fig.2.9.

The largest group is cell membrane proteins with 155 hits, which is consistent with the fact that many well characterized NUMB binding partners are membrane associated, such as NOTCH. Other top ranking groups include transcription (98 hits), kinases (69 hits), ion channels (49 hits), mitochondrial (37 hits), ubl conjugate pathway (28 hits) and protein tyrosine phosphatases (28 hits). A full list of NUMB-PTB binding candidates is attached in Suppl.S6. Aside from the tyrosine kinases that are highly enriched in this list, the PTP group is of special interest. We have shown above that NUMB-PTB binds to a number of RTKs both *in vivo* and *in vitro*, and in next chapter, a mechanistic insight into NUMB-ALK interplay in cancer will be presented. Together, these data reveal the possibility that NUMB may act as a universal binding partner and regulator of RTKs, while in tumorigenesis, NUMB may also alter the progression of RTK positive cancer types. Thus, if NUMB is involved in broad interactions with PTPs resulting in biological function, we can speculate that a regulatory mechanism in the p-Tyr-mediated signaling network may also exist, in which NUMB acts as a common node to connect RTKs and PTPs in signal regulation. A test has been performed to examine the NUMB-PTP binding. We have confirmed NUMB-PTB bind to two motifs from PTP-9 and PTP-13, both with high affinity (Fig.2.10A-B).

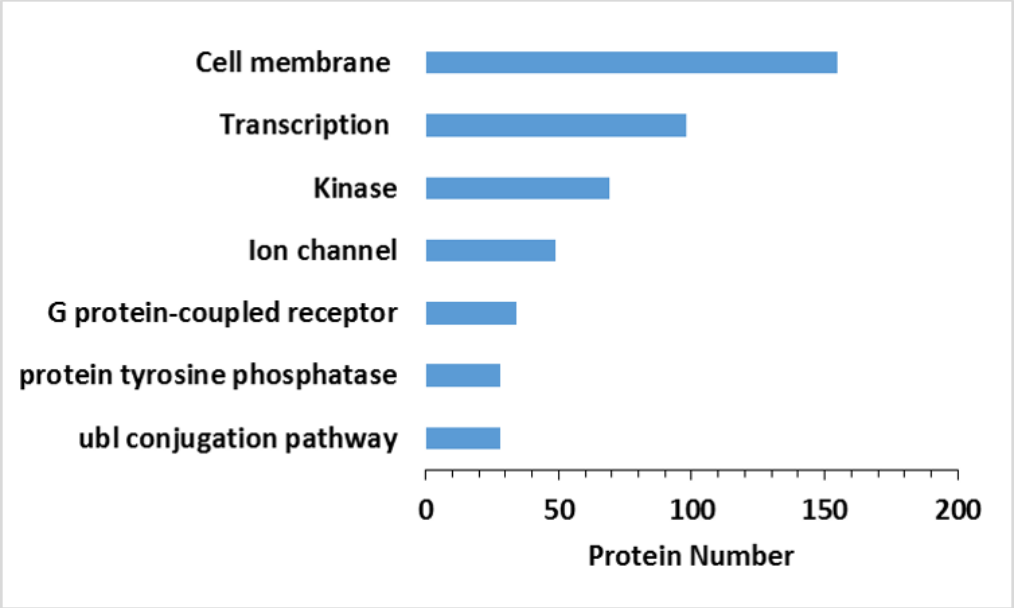


Figure.2.9 Prediction of NUMB-PTB binding motifs genome-wide

Classification of the top 1000 predicted NUMB-PTB binding motifs. These motifs were generated from 1083 proteins, and grouped according to GO annotations.

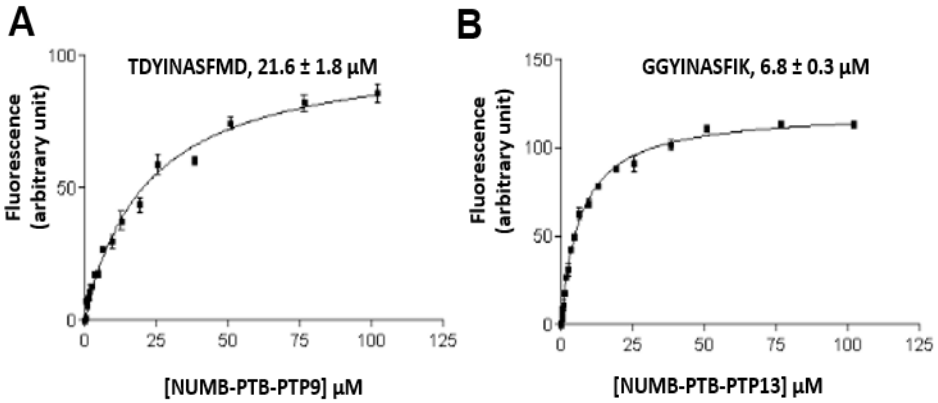


Figure.2.10 NUMB binds to motifs in PTP-9 and PTP-13

The binding curves for two motifs: TDYINASFMD from PTP-9 (A) and GGYINASFIK from PTP-13 (B). The binding affinities are 21.6 ± 1.8 µM and 6.8 ± 0.3 µM, respectively.

2.5 Discussion

2.5.1 An integrative array strategy for systematic investigation of domain-peptide interaction mediated PPI network.

Protein-protein interactions play a central role in almost all cellular activities. A number of high-throughput approaches have been well established for mapping PPI networks on a large scale. However, each approach has its own limitations. Although protein and peptide arrays may be used in systematic PPI identification *in vitro* (Jones et al., 2006; Tong et al., 2002), inconsistent data are often observed using protein or peptide arrays alone, perhaps due to the fact that some peptides may have improperly oriented spatial arrangements (for instance, the peptide is not exposed in the protein surface). In our study of the NUMB-PTB interactome, we introduced an integrative array strategy that overcomes these limitations. The high-throughput interactome screen began with a ProtoArray, which generated high-quality quantitative data representing physical interactions between NUMB-PTB and its binding partners. All potential ProtoArray hits for NUMB-PTB binding sites were individually verified in a peptide array and their binding affinities were quantitatively determined by FP. Further analysis suggested a good correlation in binding affinity quantifications between the protein and peptide arrays. In subsequent *in vivo* tests, a number of selected high scoring interactions (NUMB-RTKs) were verified and as presented in the next chapter, the NUMB-ALK interaction was shown to be biologically relevant. Although the physiological functions of the majority of interactions in this interactome await validation, our study provides an unbiased strategy to systematically explore domain-peptide mediated PPI networks.

2.5.2 The genome-wide prediction of NUMB-PTB binding partners

According to preliminary studies, NUMB-PTB was only known to bind NxxF and NxxY motifs. Here we systematically investigated NUMB-PTB domain binding specificity by determining the amino acid preferences from -5 to +2 positions (defining F/Y at zero position) via an integrative strategy combining *in vitro* and *in vivo* assays. Based on the determined level of specificity, we scanned the NUMB-PTB binding motif genome-wide

and analyzed the top 1000 candidates. It should be noted that a number of previously identified NUMB binding partners rank highly in this list, including NOTCH, cadherin, Integrin and EphB2.

Membrane associated proteins, RTKs and PTPs are of particular interest in this study and comprise the largest group in the candidate list. This is consistent with the fact that NUMB is an endocytic protein. It is well known that NUMB localizes at the cell membrane via its PTB domain and recruits other endocytic proteins. As introduced earlier, RTKs are the most common oncogenic proteins and major therapeutic targets in cancer treatment. Presented in chapter 3, the NUMB-ALK interaction is shown to have biological relevance in which NUMB regulates ALK activity antagonistically during tumorigenesis in an isoform dependent manner. Considering the general function of NUMB in endocytosis, it can be speculated that NUMB may act similarly with other RTKs identified in our study, as well as with other membrane localized proteins. If supported by further evidence, NUMB may play a role as a crucial indirect effector of trans-membrane signaling transduction via altering membrane receptor abundance. Particularly in growth factor stimulated RTK signaling, if NUMB also assists in recruiting PTPs, an extra layer of regulation may be employed to the p-Tyr signaling network to make finely tuning to p-Tyr dynamics.

Interestingly, proteins associated with transcription are also highly enriched in the prediction. Previously, few nuclear binding partners were identified for NUMB, even though NUMB can localize in the nucleus (Dhami et al., 2013; Martin-Blanco et al., 2014; Wang et al., 2015). The NUMB-p53 interaction has been thoroughly studied. Although NUMB binding has been shown to regulate the transcription of various genes downstream of p53, in this case, NUMB is only implicated in transcription regulation indirectly via p53 stabilization. It is still uncertain whether the NUMB-p53 interaction alters the p53 transcription activity at a molecular level. However, considering the complex interactome of NUMB, we cannot exclude the possibility that NUMB may be involved in transcription broadly by directly interacting with transcription associated proteins.

2.5.3 NUMB may act as a universal regulator to RTKs

Previously, NUMB has been reported to interact with three RTKs: EphB2, TrkB and EGFR, of which all three interactions have been shown biologically relevance. The NUMB-EphB2 interaction is involved in ephrin-B1-induced spine development and maturation (Nishimura et al., 2006). The NUMB-TrkB interaction alters TrkB's chemotactic response to BDNF (brain-derived neurotrophic factor) ligand (Zhou et al., 2011). The NUMB-EGFR interaction is slightly more mechanically demonstrated; it has been shown that NUMB binding decreases EGFR at the protein level which regulates cell fate determination in glioblastoma stem-like cells (Jiang et al., 2012). However, little direct evidence drawing solid conclusions for the functions of NUMB-RTK interactions has been published. Here, our study demonstrates that RTKs are highly enriched in the NUMB interactome and two typical oncogenic RTKs ALK and ErbB2 are directly interacting with NUMB *in vivo*. This discovery may bring NUMB to the center of tumorigenesis as a potential therapeutic target for RTK positive cancer types.

2.6 Reference

Arkin, M. R., and Wells, J. A. (2004). Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat Rev Drug Discov *3*, 301-317.

Berdnik, D., Torok, T., Gonzalez-Gaitan, M., and Knoblich, J. A. (2002). The endocytic protein alpha-Adaptin is required for numb-mediated asymmetric cell division in Drosophila. Developmental cell *3*, 221-231.

Brown, K. R., Otasek, D., Ali, M., McGuffin, M. J., Xie, W., Devani, B., Toch, I. L., and Jurisica, I. (2009). NAViGaTOR: Network Analysis, Visualization and Graphing Toronto. Bioinformatics *25*, 3327-3329.

Chang, D. D., Wong, C., Smith, H., and Liu, J. (1997). ICAP-1, a novel beta1 integrin cytoplasmic domain-associated protein, binds to a conserved and functionally important NPXY sequence motif of beta1 integrin. J Cell Biol *138*, 1149-1157.

Chien, C. T., Wang, S., Rothenberg, M., Jan, L. Y., and Jan, Y. N. (1998). Numb-associated kinase interacts with the phosphotyrosine binding domain of Numb and antagonizes the function of Numb in vivo. Molecular and cellular biology *18*, 598-607.

Colaluca, I. N., Tosoni, D., Nuciforo, P., Senic-Matuglia, F., Galimberti, V., Viale, G., Pece, S., and Di Fiore, P. P. (2008). NUMB controls p53 tumour suppressor activity. Nature *451*, 76-80.

Dhami, G. K., Liu, H., Galka, M., Voss, C., Wei, R., Muranko, K., Kaneko, T., Cregan, S. P., Li, L., and Li, S. S. (2013). Dynamic methylation of Numb by Set8 regulates its binding to p53 and apoptosis. Molecular cell *50*, 565-576.

Dho, S. E., French, M. B., Woods, S. A., and McGlade, C. J. (1999). Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. J Biol Chem *274*, 33097-33104.

Di Marcotullio, L., Ferretti, E., Greco, A., De Smaele, E., Po, A., Sico, M. A., Alimandi, M., Giannini, G., Maroder, M., Screpanti, I., and Gulino, A. (2006). Numb is a suppressor of Hedgehog signalling and targets Gli1 for Itch-dependent ubiquitination. Nature cell biology *8*, 1415-1423.

Filardo, E. J., Brooks, P. C., Deming, S. L., Damsky, C., and Cheresh, D. A. (1995). Requirement of the NPXY motif in the integrin beta 3 subunit cytoplasmic tail for melanoma cell migration in vitro and in vivo. J Cell Biol *130*, 441-450.

Guenette, S. Y., Chen, J., Ferland, A., Haass, C., Capell, A., and Tanzi, R. E. (1999). hFE65L influences amyloid precursor protein maturation and secretion. J Neurochem *73*, 985-993.

Gulino, A., Di Marcotullio, L., and Screpanti, I. (2010). The multiple functions of Numb. Exp Cell Res *316*, 900-906.

Guo, M., Jan, L. Y., and Jan, Y. N. (1996). Control of daughter cell fates during asymmetric division: interaction of Numb and Notch. Neuron *17*, 27-41.

Hill, K., Li, Y., Bennett, M., McKay, M., Zhu, X., Shern, J., Torre, E., Lah, J. J., Levey, A. I., and Kahn, R. A. (2003). Munc18 interacting proteins: ADP-ribosylation factor-dependent coat proteins that regulate the traffic of beta-Alzheimer's precursor protein. J Biol Chem *278*, 36032-36040.

Howell, B. W., Lanier, L. M., Frank, R., Gertler, F. B., and Cooper, J. A. (1999). The disabled 1 phosphotyrosine-binding domain binds to the internalization signals of transmembrane glycoproteins and to phospholipids. Mol Cell Biol *19*, 5179-5188.

Jiang, X., Xing, H., Kim, T. M., Jung, Y., Huang, W., Yang, H. W., Song, S., Park, P. J., Carroll, R. S., and Johnson, M. D. (2012). Numb regulates glioma stem cell fate and growth by altering epidermal growth factor receptor and Skp1-Cullin-F-box ubiquitin ligase activity. Stem cells *30*, 1313-1326.

Jones, R. B., Gordus, A., Krall, J. A., and MacBeath, G. (2006). A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature *439*, 168-174.

Juven-Gershon, T., Shifman, O., Unger, T., Elkeles, A., Haupt, Y., and Oren, M. (1998). The Mdm2 oncoprotein interacts with the cell fate regulator Numb. Molecular and cellular biology *18*, 3974.

Kaushansky, A., Allen, J. E., Gordus, A., Stiffler, M. A., Karp, E. S., Chang, B. H., and MacBeath, G. (2010). Quantifying protein-protein interactions in high throughput using protein domain microarrays. Nat Protoc *5*, 773-790.

Kinoshita, A., Whelan, C. M., Smith, C. J., Mikhailenko, I., Rebeck, G. W., Strickland, D. K., and Hyman, B. T. (2001). Demonstration by fluorescence resonance energy transfer of two sites of interaction between the low-density lipoprotein receptor-related protein and the amyloid precursor protein: Role of the intracellular adapter protein Fe65. Journal of Neuroscience *21*, 8354-8361.

Li, L., Wu, C., Huang, H., Zhang, K., Gan, J., and Li, S. S. (2008a). Prediction of phosphotyrosine signaling networks using a scoring matrix-assisted ligand identification approach. Nucleic acids research *36*, 3263-3273.

Li, L., Wu, C. G., Huang, H. M., Zhang, K. Z., Gan, J., and Li, S. S. C. (2008b). Prediction of phosphotyrosine signaling networks using a scoring matrix-assisted ligand identification approach. Nucleic acids research *36*, 3263-3273.

Li, S. C., Zwahlen, C., Vincent, S. J. F., Jane McGlade, C., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (1998). Structure of a Numb PTB domain-peptide complex suggests a basis for diverse binding specificity. Nature Structural Biology *5*, 1075-1083.

Martin-Blanco, N. M., Checquolo, S., Del Gaudio, F., Palermo, R., Franciosa, G., Di Marcotullio, L., Gulino, A., Canelles, M., and Screpanti, I. (2014). Numb-dependent integration of pre-TCR and p53 function in T-cell precursor development. Cell death & disease *5*, e1472.

McGill, M. A., Dho, S. E., Weinmaster, G., and McGlade, C. J. (2009). Numb regulates post-endocytic trafficking and degradation of Notch1. The Journal of biological chemistry *284*, 26427-26438.

McGill, M. A., and McGlade, C. J. (2003). Mammalian numb proteins promote Notch1 receptor ubiquitination and degradation of the Notch1 intracellular domain. J Biol Chem *278*, 23196-23203.

Nie, J., McGill, M. A., Dermer, M., Dho, S. E., Wolting, C. D., and McGlade, C. J. (2002). LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation. The EMBO journal *21*, 93-102.

Nishimura, T., and Kaibuchi, K. (2007). Numb controls integrin endocytosis for directional cell migration with aPKC and PAR-3. Developmental cell *13*, 15-28.

Nishimura, T., Yamaguchi, T., Tokunaga, A., Hara, A., Hamaguchi, T., Kato, K., Iwamatsu, A., Okano, H., and Kaibuchi, K. (2006). Role of numb in dendritic spine development with a Cdc42 GEF intersectin and EphB2. Molecular biology of the cell *17*, 1273-1285.

Ostman, A., Hellberg, C., and Bohmer, F. D. (2006). Protein-tyrosine phosphatases and cancer. Nature reviews Cancer *6*, 307-320.

Pece, S., Confalonieri, S., P, R. R., and Di Fiore, P. P. (2011). NUMB-ing down cancer by more than just a NOTCH. Biochim Biophys Acta *1815*, 26-43.

Pece, S., Serresi, M., Santolini, E., Capra, M., Hulleman, E., Galimberti, V., Zurrida, S., Maisonneuve, P., Viale, G., and Di Fiore, P. P. (2004). Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. The Journal of cell biology *167*, 215-221.

Rhyu, M. S., Jan, L. Y., and Jan, Y. N. (1994). Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. Cell *76*, 477-491.

Rodriguez, M., Li, S. S., Harper, J. W., and Songyang, Z. (2004). An oriented peptide array library (OPAL) strategy to study protein-protein interactions. J Biol Chem *279*, 8802-8807.

Salcini, A. E., Confalonieri, S., Doria, M., Santolini, E., Tassi, E., Minenkova, O., Cesareni, G., Pelicci, P. G., and Di Fiore, P. P. (1997). Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. Genes Dev *11*, 2239-2249.

Sato, K., Watanabe, T., Wang, S., Kakeno, M., Matsuzawa, K., Matsui, T., Yokoi, K., Murase, K., Sugiyama, I., Ozawa, M., and Kaibuchi, K. (2011). Numb controls E-cadherin endocytosis through p120 catenin with aPKC. Molecular biology of the cell *22*, 3103-3119.

Smith, C. A., Lau, K. M., Rahmani, Z., Dho, S. E., Brothers, G., She, Y. M., Berry, D. M., Bonneil, E., Thibault, P., Schweisguth, F., *et al.* (2007). aPKC-mediated phosphorylation regulates asymmetric membrane localization of the cell fate determinant Numb. The EMBO journal *26*, 468-480.

Susini, L., Passer, B. J., Amzallag-Elbaz, N., Juven-Gershon, T., Prieur, S., Privat, N., Tuynder, M., Gendron, M. C., Israel, A., Amson, R., *et al.* (2001). Siah-1 binds and regulates the function of Numb. Proceedings of the National Academy of Sciences of the United States of America *98*, 15067-15072.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., *et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic acids research *43*, D447-452.

Tokumitsu, H., Hatano, N., Inuzuka, H., Sueyoshi, Y., Yokokura, S., Ichimura, T., Nozaki, N., and Kobayashi, R. (2005). Phosphorylation of Numb family proteins. Possible involvement of Ca2+/calmodulin-dependent protein kinases. J Biol Chem *280*, 35108-35118.

Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., *et al.* (2002). A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. Science *295*, 321-324.

Uemura, T., Shepherd, S., Ackerman, L., Jan, L. Y., and Jan, Y. N. (1989). numb, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos. Cell *58*, 349-360.

Uhlik, M. T., Temple, B., Bencharit, S., Kimple, A. J., Siderovski, D. P., and Johnson, G. L. (2005). Structural and evolutionary division of phosphotyrosine binding (PTB) domains. Journal of molecular biology *345*, 1-20.

Wang, C., Cui, T., Feng, W., Li, H., and Hu, L. (2015). Role of Numb expression and nuclear translocation in endometrial cancer. Oncology letters *9*, 1531-1536.

Wang, Z., Sandiford, S., Wu, C., and Li, S. S. (2009). Numb regulates cell-cell adhesion and polarity in response to tyrosine kinase signalling. The EMBO journal *28*, 2360-2373.

Webb, B., and Sali, A. (2014). Protein structure modeling with MODELLER. Methods in molecular biology *1137*, 1-15.

Zhong, W., Jiang, M. M., Schonemann, M. D., Meneses, J. J., Pedersen, R. A., Jan, L. Y., and Jan, Y. N. (2000). Mouse numb is an essential gene involved in cortical neurogenesis.

Proceedings of the National Academy of Sciences of the United States of America *97*, 6844-6849.

Zhou, P., Alfaro, J., Chang, E. H., Zhao, X., Porcionatto, M., and Segal, R. A. (2011). Numb links extracellular cues to intracellular polarity machinery to promote chemotaxis. Dev Cell *20*, 610-622.

Zwahlen, C., Li, S. C., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (2000). Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. The EMBO journal *19*, 1505-1515.

Chapter 3

3 NUMB regulates ALK endocytosis and activity in an isoform-dependent manner

3.1 Abstract

NUMB is an evolutionarily conserved protein that plays a multifaceted role in cellular homeostasis and is implicated in a variety of cellular events including cell adhesion and migration, endocytosis, polarity and cell fate determination. In cancer, NUMB is conventionally defined as a tumor suppressor, as loss of NUMB usually results in a disorder of these physiological processes. Conversely, we find that NUMB also promotes tumorigenesis via directing differentiated post-endocytic destinations of receptor tyrosine kinase ALK in an isoform dependent manner. NUMB directly interacts with ALK and mediates ALK endocytosis. After internalization, p66-NUMB mainly directs ALK to the lysosomal degradation pathway, thus attenuating the overall ALK level and the downstream MAPK proliferation signaling. Antagonistically, p72-NUMB promotes ALK recycling to the cell membrane and maintains the cell in an ALK active status. Our study explains some controversial findings about NUMB functions in tumorigenesis and reveals the tumorigenic potential of the p72-NUMB isoform, and also provides mechanistic insight into ALK regulation.

3.2 Introduction

Anaplastic lymphoma kinase (ALK) is a member of the orphan receptor tyrosine kinase (RTK) subfamily, and originally was described as a fusion protein of the t (2;5) (p23;q35) translocation expressed in Anaplastic Large Cell Lymphomas (ALCL) (Morris et al., 1994; Shiota et al., 1994). Normally, the ALK receptor is only transiently expressed in the nervous system during early developmental stages and maintained at low level in adult tissue, with the exception of some neuronal cells (Morris et al., 1994). This gene is considered non-essential for development as only a mild behavioral phenotype was observed in knockout mice (Bilsland et al., 2008). However, deregulation of ALK activity, including mutations, amplifications and translocations, have been characterized in a subset of cancers. For instance, two ALK constitutively-active mutants F1174L and R1275Q are frequently found in neuroblastomas, which account for 15% of all childhood cancer deaths (George et al., 2008). ALK amplification has also been detected in neuroblastoma, colorectal cancer and non-small cell lung cancer (NSCLC) (Bavi et al., 2013; Salido et al., 2011; Wang et al., 2013). It should be noted that approximately 4%~7% of NSCLC involve an EML4-ALK fusion kinase that is self-activated and shown to promote and maintain the malignant behavior of NSCLC cells (Sasaki et al., 2010; Soda et al., 2007). Similar to other RTK positive tumors, the deregulated ALK promotes tumorigenesis by constitutively stimulating various cancer-related signaling pathways including the MAPK proliferation pathway and the PI3K survival pathway. Since ALK levels are low in healthy adult tissues, selective ALK inhibition tends to result in less adverse effects without compromising efficacy (Crescenzo and Inghirami, 2015). The FDA granted accelerated approvals in 2011 and 2014 for two generations of ALK inhibitors (crizotinib and ceritinib) in treating ALK positive metastatic NSCLC. In addition, another inhibitor (alectinib) was recently approved for NSCLC patients who could not tolerate crizotinib treatment.

Despite the oncogenic potential and therapeutic value of ALK, the mechanism to ALK regulation remains elusive. Here we report that NUMB, a well-known adaptor protein, directly interacts with ALK, and two NUMB isoforms (p72 and p66) regulate ALK activity antagonistically by directing ALK to different post-endocytic pathways. NUMB is a membrane-associated protein that is highly conserved from fly to human and consistently

plays an essential role in the maintenance of cellular homeostasis in the nervous system (Dho et al., 1999; Pece et al., 2011). The NUMB gene was the first isolated cell fate determinant from *Drosophila*, whereby asymmetric distribution of NUMB during mitosis produced diverse neural cell fates (Uemura et al., 1989). In tumorigenesis, NUMB functions as an adaptor, and its asymmetric distribution has been implicated in a variety of cancer-related events and signaling pathways as well (Gulino et al., 2010; Pece et al., 2011). Conventionally, NUMB is defined as a tumor suppressor due to its negative regulation of the Notch signaling pathway and cell proliferation by promoting Notch endocytic degradation (Couturier et al., 2012; McGill et al., 2009; McGill and McGlade, 2003), or by preventing MDM2-mediated ubiquitination of p53 (Colaluca et al., 2008). Nevertheless, NUMB is also observed to promote tumorigenesis in an alternative splicing (AS) dependent manner.

The NUMB gene undergoes alternative splicing to produce multiple isoforms. The four isoforms well characterized in both *Drosophila* and human share a similar structure and contain an N-terminal phosphotyrosine binding (PTB) domain, a proline rich region (PRR), two Asp-Pro-Phe (DPFs) motifs and one Asn-Pro-Phe (NPF) motif at the C-terminal. The isoforms are differentiated by the inclusion or exclusion of exon 3 in the PTB domain or exon 9 in the PRR (Bork and Margolis, 1995; Dho et al., 1999; Salcini et al., 1997; Verdi et al., 1996). The four isoforms exhibited distinct functions in *Drosophila* neuronal development (Verdi et al., 1999). Two PPR^{S} isoforms primarily stimulate differentiation while two PRR^{L} isoforms promote proliferation. Consistently, PRR^{L} isoforms have been characterized in multiple cancer types, including breast cancer, colon cancer and lung cancer (Misquitta-Ali et al., 2011). It has also been reported that RBM5/6 and RBM10 splicing factors, which are responsible for the NUMB exon 9 inclusion or exclusion, regulate cancer cell proliferation conversely (Bechara et al., 2013). The mechanism of these controversial effects is not well elucidated, but it is speculated that the exon 9 inclusion isoforms p72 and p71 may act antagonistically to exon 9 skipping isoforms p66 and p65 in cell proliferation. Indeed, in some lung cancer cells it has been shown that inclusion of exon 9 is correlated with increased cell proliferation (Bechara et al., 2013; Westhoff et al., 2009).

In this study, we thoroughly studied the interaction between ALK and NUMB, which was identified from a preliminary high throughput screening of NUMB interactors, as presented in Chapter 2. Our results reveal that the NUMB-PTB domain binds to N1477MAF1480 and N1585YGY1586 motifs in ALK *in vivo*, and the interaction directly regulates ALK internalization. In post-endocytic trafficking, the p66-NUMB isoform mainly promotes ALK degradation via the lysosomal pathway thus decreasing the activity of ALK and downstream MAPK proliferation signaling. Conversely, the p72-NUMB isoform recycles ALK back to the cell membrane to maintain ALK activity in cells. Our findings provide mechanistic insight into ALK regulation and the tumorigenic potential of the p72-NUMB isoform, and further reveal markers of potential clinical value.

3.3 Material and methods

3.3.1 Cell culture

IMR-5 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 µg/ml penicillin/streptomycin and L-glutamine. HEK293 cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% FBS, 100 µg/ml penicillin/streptomycin and L-glutamine. Cells were incubated at 37 °C in a humidified atmosphere containing 5% carbon dioxide.

3.3.2 Transient transfection

The plasmid DNA was first diluted in serum-free Minimum Essential Medium (MEM) to a final concentration 0.01 µg/µl. For each 10-cm dish with 10ml culture medium, 40 µl of X-tremeGENE HP DNA transfection reagent (Roche) was mixed with 1ml plasmid DNA in MEM. The mixture was added to the culture dish after 15 min incubation at room temperature. Following transfection, cells were incubated for 48~72 hours to allow the ectopic gene expression.

3.3.3 siRNA interference

siRNA was synthesized to target either p66-NUMB, p72 NUMB or NUMB. For each 35-mm dish, siRNA was first diluted with MEM to a final concentration of 2.5 µM in a 200 µl total volume. 5 µl DharmaFECT reagent (GE) was added to 195 µl serum-free MEM and gently mixed. The siRNA and reagent were then mixed and added to the culture dish with 1.6 ml culture medium after 15 min incubation at room temperature. Cells were incubated for 48~72 hours before further analysis.

3.3.4 Immunoprecipitation and western blotting

The cultured cells were lysed in cold mammalian cell lysis buffer (1% Triton X-100, 50 mM Tris-pH 7.2, 150 mM NaCl, 2 mM MgCl2, 0.1 mM EDTA, 0.1 mM EGTA, 0.5 mM DTT, 10 mM NaF) containing complete protease inhibitors (Roche). Cell debris was

removed by centrifugation, and 500 µg of supernatant protein was incubated for four hours in the presence of 1 µg antibody and 30 µl of 50% slurry protein G-Sepharose beads (Roche) at 4°C. The beads were subsequently washed three times with lysis buffer and boiled with SDS-loading dye. The immunoprecipitates were resolved by SDS-PAGE. All proteins were then transferred to PVDF membrane by semi-dry transferring method (Bio-Rad), detected by immunoblotting with appropriate antibodies, and visualized by Enhanced Chemiluminescence (ECL).

3.3.5 Recombinant protein expression and GST pull-down

The BL21 strain of *Escherichia coli* was transformed with pGEX6P3-NUMB (PTB^L/PTB^S). Positive colonies were shaked in Lysogeny Broth (LB) medium at 37°C to reach OD_{600} 0.6~0.8, then the cultures were induced with 0.5 mM IPTG for 16 hours at 18°C. The bacterial cells were harvested and pellets were suspended in PBS buffer containing complete protease inhibitors (Roche). Triton X-100 was added to a final concentration of 2%, lysozyme was added to a final concentration 1mg/ml and benzonase was added to a final concentration 20 units/ml. The suspension was sonicated six times (10 seconds each) on ice and then lysed for 30 min at room temperature. The lysate was centrifuged at 15,000 × g for 30 min at 4°C for collecting the supernatant. Purification of GST-tagged proteins was performed with glutathione resin (GE healthcare). The resin was washed with PBS buffer three times, and then the lysate was loaded to the resin followed by 3 times washing with PBS buffer. To determine the purity of GST-proteins, a small amount of resin was boiled with SDS-loading dye and then analyzed by SDS-PAGE/Coomassie staining. Pull-down was performed by incubating mammalian whole cell lysate with GST-NUMB PTB^L, GST-NUMB PTB^S or GST beads for 4 hours at 4°C, followed by 3 washes with mammalian cell lysis buffer. Proteins that bound to GST or GST-PTB were resolved on SDS-PAGE and identified by Western Blotting.

3.3.6 Gene truncation and site-directed mutagenesis

Truncated forms of NUMB were constructed by standard PCR and re-ligation procedures. Five gene mutants, p66-NUMB-F162V, p72-NUMB-F162V, ALK-N1477A, ALK-

N1583A, ALK-1477A/1583A were constructed by site-directed mutagenesis. A pair of complementary mutagenic primers were designed for individual mutants. The entire plasmids were amplified using high fidelity DNA polymerase. Then Dpn1 enzyme was used to digest the methylated template plasmid. Undigested plasmids were recycled and used for transformation. Positive colonies were identified by gene sequencing.

3.3.7 Immunostaining and microscopic quantification

Cell were cultured in 35-mm glass-bottomed dishes (P35G-1.0C, Matek) under standard conditions. When cells grew to the 20%~40% confluence, they were fixed with 4% paraformaldehyde and permeabilized with 0.2% triton X-100 for 10~30 mins. For staining, the cells were rinsed with PBS buffer three times and incubated with 3% BSA in PBS to block non-specific signal. The primary antibodies were diluted with 3% BSA in PBS at 1:50~ 1:200 ratios and added to the dishes for overnight incubation. Then the cells were rinsed with PBS buffer three times and incubated with secondary fluorescent antibodies (Invitrogen, 1:1000 dilution) for 1 hour. Finally, the cells were rinsed with PBS buffer three times and covered with mounting medium containing DAPI (4',6-diamidino-2-phenylindole) fluorescent dye. For microscopic observation, a Zeiss Meta 510 LSM was used. To standardize images captured, all fluorescent images from different samples were obtained under the same setting of microscope. Z-stack scanning was applied to all samples with the same step distance.

All images were processed and analyzed by ImageJ software (imagej.nih.gov/ij/). Image exhibiting co-localization were merged from the same z-stack slices in two different channels. For co-localization quantification, the Pearson product-moment correlation (Pearson's) was quantified by JACoP (Just Another Colocalization Plugin) plugin of ImageJ software. From the unprocessed images, individual cells were picked by ROI (Regions of Interest) manager plugin and their Pearson's coefficient values were determined by JACop plugin under standard settings.

3.3.8 Peptide synthesis

Both free and membrane-bound peptides were synthesized using an automatic Intavis AG workstation from amino acid monomers protected with Fmoc (9-fluorenylmethyl-oxycarbonyl). Rink-resin (Rink-NH_2) was used to couple the first amino acid in free peptide synthesis, whereas the amine-derivate cellulose membrane (cellulose-NH_2) was made to couple the first amino acid in on-membrane peptide synthesis. In each synthesis cycle, the carboxyl group of Fmoc-protected amino acid (Fmoc-R-COOH) was first linked to the amine group of the previous amino acid (or Rink-NH_2) through an amide bond. All unoccupied amine groups were then blocked (acetylated) by acetic anhydride to prevent incorrect amide bonds forming in subsequent cycles. Next, the Fmoc group was removed by piperidine (de-protecting) to release the free amine group primed for linking the carboxyl group of next amino acid residue. Fluorescein-NHS was linked to the amine group of the last amino acid in free peptide synthesis. After synthesis, the on-membrane peptides were treated with a mixture containing 47.5% TFA (trifluoroacetic acid), 1.5% TIPS (tri-isopropylsilane) and 51% water, to remove all other protecting groups on amino acid side chains. The free peptides were treated with a mixture containing 95% TFA, 3% TIPS and 2% water to de-protect side chains while cleaving the peptide from Rink-resin.

3.3.9 Determination of dissociation constant

The dissociation constant between protein and peptide was determined by fluorescence polarization (FP) assay. The proteins and fluorescein labelled peptides were prepared following the procedures introduced above. To reduce variation, two independent tests were prepared simultaneously. 384-well flat bottom plates (Corning-3575) were used for sample reading and PerkinElmer Envision 2103 plate reader was used for evaluating FP. The peptides (approximately 2 μM each) were dissolved in 100 μl DMSO and 3 μl of this was diluted 100-fold in water. The concentrations of the proteins were determined and adjusted to 50 μM. 2-fold serial dilutions were setup to create 16 protein concentrations ranging from 0 to 50 μM. 30μl of each protein sample was mixed with 5 μl peptide in the 384-well plate and the fluorescence polarization was determined by the plate reader. We

use an approximate equation for non-linear fitting, by assuming the concentration of peptide was much lower than KD value:

$\Delta FP = FPobs - FP0 = FPmax \times [domain] / (KD + [domain])$

The full equation without the approximation can be found in this study (Kaushansky et al., 2010).

3.3.10 Cell surface biotinylation assay and recycling assay

The surface protein biotinylation based assays were performed to measure the internalization and recycling of ALK proteins. The HEK293/ALK cells were treated as described in Fig.3.3 and Fig.3.4. Briefly, ALK on cell membrane can be biotinylated at 4°C. Returning the cells to 37°C, the biotin-ALK will eventually internalize and undergo endosomal sorting, which directs biotin-ALK to either the recycling pathway or degradation pathway. If the cells are transferred to 18°C incubation, internalization will still take place but the recycling or degradation will be slow or stopped. These different incubation conditions, combined with MESNA (sodium 2-mercaptoethanesulfonate) or GSH surface biotin stripping as well as cell fractionation, will generate a sample set for evaluating the destination and localization of biotinylated ALK at different time points. A detailed description of this method can be found in this study (Smith et al., 2004)

3.3.11 Soft agar assay

Media with the agar concentrations of 1% and 0.7% were first prepared and autoclaved. For the base layer, 1% agar was melted and cooled to 40°C in a water bath. Meanwhile, 2×RPMI with 20% FBS and 200 μg/ml penicillin/streptomycin was warmed in a 40°C water bath. Next, equal volumes to the two solutions were mixed to result in 0.5% agar, 1×RPMI and 10% FBS mixture, and 1.5 ml of the mixture was quickly added to each well of a 6-well plate. For the top layer, 0.7% agar was used for making the mixture. Cells were trypsinized and the cell density was adjusted to 200,000 cells/ml in warm PBS buffer. 30,000 cells were suspended gently in 9 ml mixture containing 0.35% agar, 1×RPMI and 10% FBS. 1.5 ml of the mixture was quickly added to each well of 6-well plate on top of

the base layer. Then the plates were incubated at 37 °C in a humidified atmosphere containing 5% carbon dioxide, fed with 1×RPMI medium containing 10% FBS, 100 µg/ml penicillin/streptomycin and L-glutamine. In 10 to 30 days, when the colonies grew to proper sizes, the plates were stained with 0.005% crystal violet for 1 hour. Then the colonies were counted under stereomicroscope.

3.3.12 Cell proliferation assay

Overall cell viability was evaluated using WST-8 ((2-(2-methoxy-4-nitrophenyl)-3-(4-nitrophenyl)-5-(2,4-disulfophenyl)-2H-tetrazolium, monosodium salt). After incubation with the cells, WST-8 produced a water-soluble formazan dye in the presence of cellular dehydrogenase. Cells were cultured in 100 µl medium volume within 96-well plates. 10 µl of WST-8 solution was directly added to 100 µl medium and cells were returned to the incubation chamber and incubated for another 30 min to 1 hour. Then the 96-well plates were read for absorbance at 460nm which was proportional to the total number of viable cells.

3.4 Results

3.4.1 Identification of ALK as a novel NUMB-binding protein

A critical region in the NUMB protein is the N-terminal PTB domain, which binds to a variety of proteins that are directly involved in NUMB functions (Gulino et al., 2010). Preliminarily, we isolated several PTB binding receptor tyrosine kinases from a high-throughput screening of PTB binding partners, including ALK (Chapter 2). The endogenous NUMB-ALK interaction was first verified by co-immunoprecipitation (Co-IP) in the IMR-5 cell line (Fig.3.1A), which is a neuroblastoma cell line known for its wild type ALK expression (George et al., 2008; Tumilowicz et al., 1970). In addition, by immunostaining we found that NUMB partially co-localized with ALK in IMR-5 cells. The co-localization signal was presented in a dot-like pattern and was mainly located in cell membrane peripheral region (Fig.3.1B), suggesting NUMB-ALK interaction may occur in endocytic vesicles.

Because few cell lines were well characterized for the expression of both ALK and NUMB (isoforms) and the IMR-5 cell line was difficult to be transfected efficiently, in all following studies, we used a HEK293/ALK cell line that stably expressed wild type ALK(Moog-Lutz et al., 2005). Using this HEK293/ALK cell line, we validated the NUMB-ALK interaction by Co-IP (Fig.3.1C), then we used GST-fused PTB domains with or without exon 3 (GST-PTB^{L}, long form and GST-PTB^{S}, short form) to pull down ALK from the same cell line. Both GST fusion proteins, but not GST itself were capable of binding ALK (Fig.3.1D). To investigate whether the PTB domain is sufficient for ALK binding, we created a NUMB mutant bearing the F162V mutation within the PTB domain, which was reported to render the PTB domain incapable of motif-binding (Zwahlen et al., 2000). Upon expression of either NUMB wildtype or the F162V mutant, a marked reduction of binding to ALK was detected for F162V mutant (Fig.3.1E), indicating that the PTB domain played a critical role in NUMB-ALK binding. In addition, when a series of truncated NUMB proteins with a FLAG tag were generated (Fig.3.1F) and expressed in the HEK293/ALK cell, only the PTB-containing fragments were found to interact with ALK (Fig.3.1G).

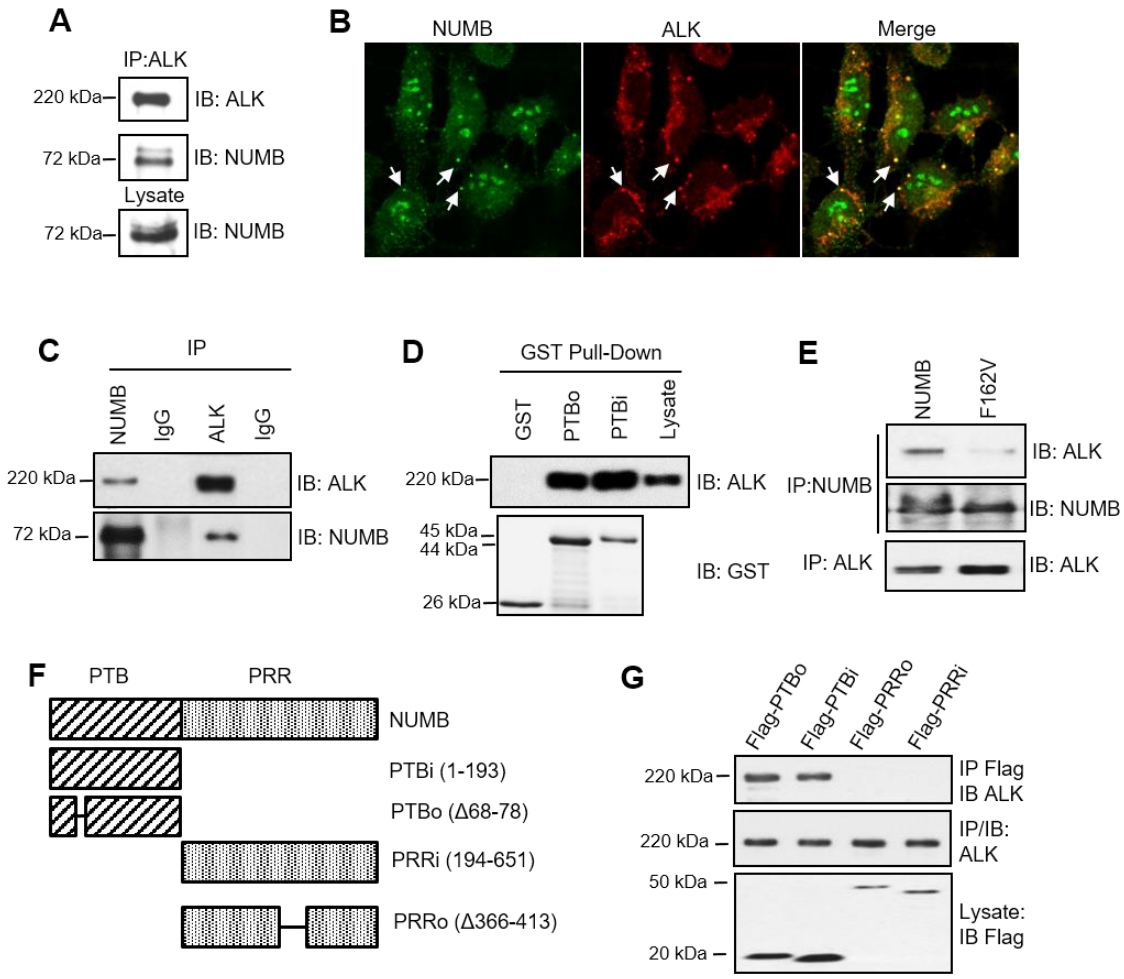


Figure.3.1 Identification of the interaction between NUMB and ALK

(A) Endogenous NUMB and ALK interaction in IMR-5 neuroblastoma cells; (B) NUMB and ALK co-localization in IMR-5 cells immuno-stained with NUMB (green) and ALK (red) antibodies. The co-localization signals were mainly present in the cell membrane periphery, and in dot-like formations, suggesting NUMB may interact with ALK in endocytic vesicles as well (arrows); (C) Endogenous NUMB interaction with ectopically-expressed ALK in HEK293/ALK stable cell line. The identical data was presented in Fig.2.8C; (D-G) The NUMB PTB domains were crucial for NUMB-ALK interactions. NUMB has two forms of PTB domains, PTBi (PTB^{L}) and PTBo (PTB^{S}) due to alternative splicing. Both GST-fused PTBi/PTBo proteins, but not GST protein, were able to pull down ALK from the HEK293/ALK cell lysate (D), and the NUMB-ALK interaction was greatly attenuated with the PTB binding-deficient mutant F162V (E), indicating that the PTB domain was responsible for ALK binding. A variety of Flag-tagged truncated forms of NUMB (F) were expressed in HERK293/ALK cells, only the PTB containing fragments were found to pulldown ALK (G).

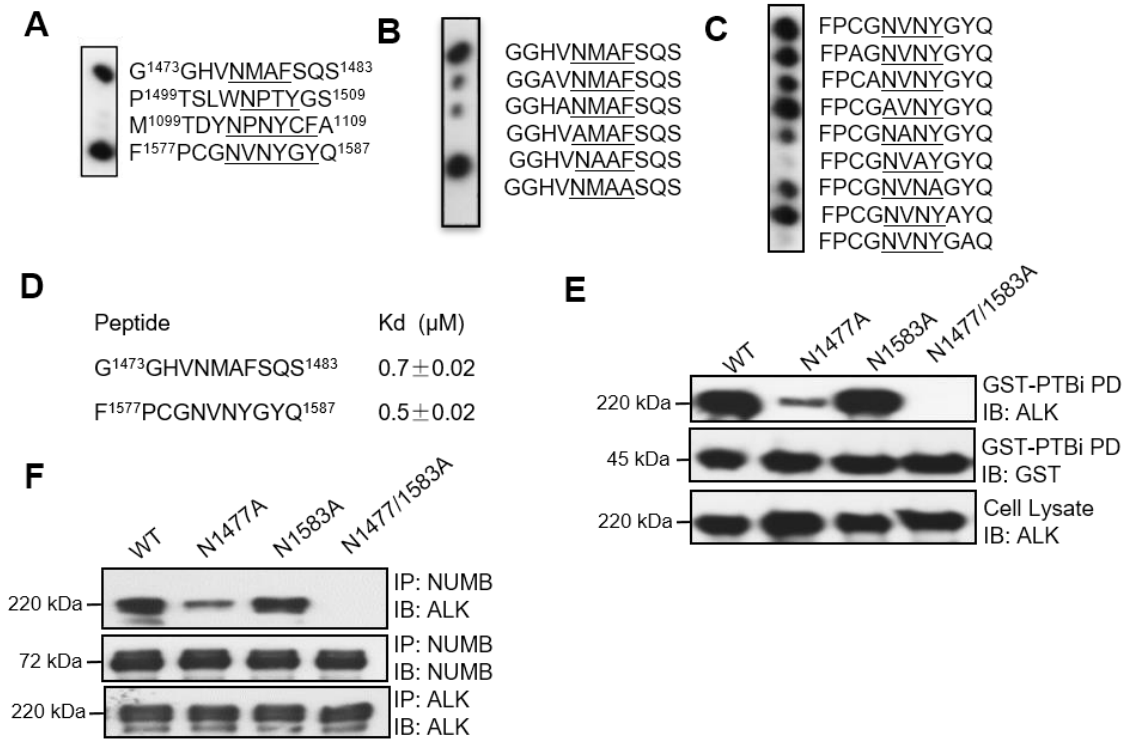


Figure.3.2 NUMB binds two motifs in ALK

(A) NUMB interaction with two NxxF/Y motifs in ALK. Four predicted NxxF/Y motifs, G[1473]GHVNMAFSQ[1482], P[1499]TSLWNPTYGS[1509], M[1099]TDYNPNYCFA[1109], and F[1577]PCGNVNYGYQ[1587], were synthesized on a cellulose membrane. Only G[1473]GHVNMAFSQ[1482] and F[1577]PCGNVNYGYQ[1587] interacted with GST-PTBi in a far western assay; (B-C) The N and F or Y residues in G[1473]GHVNMAFSQ[1482] and F[1577]PCGNVNYGYQ[1587] play a critical role in NUMB-ALK interaction. Serial alanine-scanning substitution peptides were synthesized on a cellulose membrane for G[1473]GHVNMAFSQ[1482] (B) and F[1577]PCGNVNYGYQ[1587] (C) motifs, and then blotted with GST-PTBi in a far western assay; (D) G[1473]GHVNMAFSQ[1482] and F[1577]PCGNVNYGYQ[1587] motifs exhibited high binding affinities to PTBi *in vitro*. The corresponding dissociation constant values were 0.7±0.02 μM and 0.5±0.02 μM respectively, n=3; (E-F) ALK mutations in the two motifs, N1477A and N1583A, together abolished NUMB-ALK interaction in *vivo*. In the GST-PTBi pulldown assay (E), the single mutant N1477A greatly attenuated NUMB-ALK interaction while N1583A had little effect, however, the double mutant completely abolished the interaction. In a complementary test of NUMB co-immunoprecipitation (F), an identical result was observed. These data demonstrated that both motifs were responsible for NUMB-ALK interaction with G[1473]GHVNMAFSQ[1482] playing a primary role.

The NUMB PTB domain is known to bind to NxxF/NxxY motifs (Li et al., 1998). Four such motifs were identified in the ALK intracellular region through a full sequence scanning, which were then tested for NUMB-PTB binding in an *in vitro* peptide-domain binding assay. The corresponding peptides, G[1473]GHVNMAFSQ[1482],

P^{1499}TSLW<u>NPTY</u>GS^{1509}, M^{1099}TDY<u>NPNY</u>CFA^{1109}, and F^{1577}PCGNV<u>NYGY</u>Q^{1587}, were synthesized on a cellulose membrane and then blotted with GST- PTB^L in a far-western assay. Only G^{1473}GHVNMAFSQ^{1482} and F^{1577}PCGNVNYGYQ^{1587} were found to bind to the PTB domain (Fig.3.2A). The two motifs were then subjected to alanine-scanning substitution and the resulting peptide arrays were blotted with GST- PTB^L as well (Fig.3.2B-C). The result further confirmed the importance of the N^{1477}MAF and N^{1583}YGY motifs in PTB binding and the critical role of the N and F or Y residues within in the motifs (Wang et al., 2009). Next, the binding affinities between the two motifs and PTB^L domain were determined by fluorescence polarization assay. The dissociation constants between fluorescein-labeled G^{1473}GHV<u>NMAF</u>SQ^{1482} / F^{1577}PCGNV<u>NYGY</u>Q^{1587} peptides and PTB^L were 0.7 μM and 0.5 μM respectively (Fig.3.2D). In addition, to verify this binding *in vivo*, we created ALK mutants in which either or both of Asn^{1477} and Asn^{1583} were mutated to Ala (N1477A/N1583A). These mutants were expressed in HEK293 cells transiently and the respective cell lysates were pulled down by GST-NUMB- PTB^L. As shown in Fig.3.2E, the ALK single mutant N1477A significantly weakened NUMB-ALK binding, while the N1477/N1583 double mutant completely abolished the interaction. In a complementary test of NUMB Co-IP from the same cells, an identical result was observed (Fig.3.2F).

Collectively, these data demonstrate that the PTB domain in NUMB and the N^{1477}MAF and N^{1577}YGY motifs in ALK mediate the protein-protein binding via direct physical interaction.

3.4.2 NUMB regulates ALK endocytosis

Because NUMB has been shown to be involved in the endocytosis of a number of membrane proteins (Gulino et al., 2010), we hypothesized that NUMB may play a similar role in regulating ALK endocytosis. To test this hypothesis, we examined ALK endocytosis by cell surface biotinylation and internalization assay. Specifically, in HEK293/ALK cells, membrane-localized HA-ALK was labelled with NHS-SS-biotin and then the cells were cultured for up to 60 min under normal conditions to allow internalization to occur, during which cytoplasmic biotinylated HA-ALK was pulled down by streptavidin beads at

different time points. The samples were then probed with HA antibody to determine the relative levels of internalized ALK in the cytoplasm. We observed that ALK internalized from the beginning and appeared to plateau at approximately 30 min (Fig.3.3A). To further investigate the possible function of NUMB in this process, we depleted total NUMB levels by RNA interference in HEK293/ALK cells followed by surface biotinylation and internalization assays. The NUMB knockdown cells exhibited significantly less internalized biotinylated ALK in the cytoplasm at the same time points compared to that in scramble siRNA treated cells (Fig.3.3B). Quantitatively, approximately 18% of surface labeled ALK was internalized after 60 min in NUMB knockdown cells while this number was around 34% in control cells (Fig.3.3C).

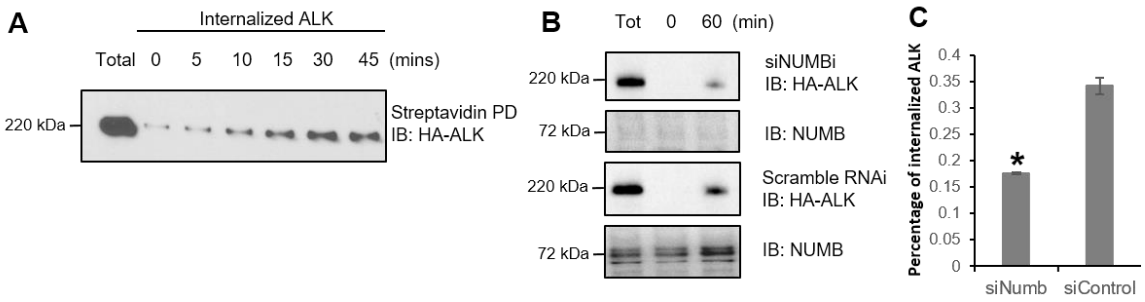


Figure.3.3 NUMB promotes ALK internalization

(A) ALK was internalized from the membrane to cytosol, as exhibited in the biotinylation and internalization assay. In HEK293/ALK cells, membrane localized HA-ALK was labelled with NHS-SS-biotin, then cytoplasmic HA-ALK levels were analyzed at different time points by streptavidin pulldown combined with HA immunoblotting; (B-C) NUMB knockdown by RNA interference significantly attenuated ALK internalization. Quantitatively, approximately 34% of surface-labelled ALK was internalized after 60 minutes in cells transfected by scrambled RNA, while the number was 18% in NUMB knockdown cells, $p<0.01$, $n=3$.

3.4.3 NUMB regulates ALK post-endocytic trafficking

The function of NUMB in endocytosis was previously elucidated by studying NUMB-regulated NOTCH receptor internalization (McGill et al., 2009). The NUMB p66 isoform was found to promote NOTCH endocytosis and lysosomal degradation which in turn inhibited NOTCH-stimulated cell proliferation signaling. However, this result is

insufficient to explain the findings that alternative NUMB splicing could increase the activity of NOTCH and its downstream signaling (Bechara et al., 2013; Westhoff et al., 2009), or that alternative splicing of NUMB is correlated with promoted cell proliferation in multiple cancer types (Misquitta-Ali et al., 2011). Mammalian NUMB has four isoforms generated by alternative splicing. The p72 isoform contains both exon 3 and exon 9, the p71 isoform contains only exon 9, the p66 isoform contains only exon 3 and the p65 isoform doesn't contain either. Since all isoforms (PTB^L/ PTB^S, or PTBi/PTBo) bind to ALK (Fig.3.1B/E), we speculate that p72/p71 may play an opposing role in post-endocytic trafficking of receptor compared to p66/p65. Instead of promoting receptor degradation, P72/p71 may recycle the receptor back to cell membrane to promote the receptor activity, which is also a default post-endocytic destination of internalized membrane receptors (Goh and Sorkin, 2013). Based on a previous study in the subcellular localizations of NUMB isoforms, p72 and p66 tend to localize to cell membrane periphery, while p71 and p65 typically localize in the cytoplasm (Dho et al., 1999). Thus, we picked isoforms p72 and p66 which generally interact with membrane localized receptors in the following endocytosis study.

The p72 and p66 isoforms were transiently expressed with a FLAG tag in HEK-293/ALK cells, and the cell surface biotinylation assay was carried out 48 hours after transfection. Next, the total biotinylated protein was pulled down with streptavidin beads 1 hour after biotinylation, and the samples were probed with FLAG antibody. The p66 isoform was found to significantly decrease the total biotinylated ALK level in cytoplasm while the p72 isoform exhibited little effect compared to the control, indicating that p66 promoted post-endocytic ALK degradation. In addition, the PTB binding deficient p66-F162V mutant and a deletion form of p66 lacking C-terminal endocytosis motif failed to affect biotinylated ALK levels, revealing their essential functions in this process (Fig.3.4A).

The decrease in biotinylated cell surface ALK caused by p66 isoform could be due to proteasome or lysosome mediated degradations. In order to identify the degradation pathway, we treated cells expressing p72 or p66 isoforms with the lysosome inhibitor concanamycin A (ConA) or the proteasome inhibitor MG132 prior to the endocytosis assay. Compared to the vehicle DMSO, MG132 treatment led to a general increase in

biotinylated ALK in all treated cells regardless whether they expressed the p72 or p66 isoforms of NUMB. In contrast, ConA treatment led to a specific increase in biotinylated ALK in cells expressing the p66 isoform (Fig.3.4B). These data indicate that the p66, but not the p72 isoform of NUMB promote the degradation of endocytic ALK via the lysosomal pathway.

To further investigate the potential function of NUMB in ALK recycling, we performed the biotinylation and recycling assays in HEK293/ALK cells transiently expressing p72 and p66 isoforms. The cells were first surface-biotinylated and incubated at 18ºC for 2 hours to allow internalization but not recycling of biotinylated ALK. Next, the cells were treated with glutathione (GSH) to remove biotin from cell surface. Following surface biotin removal, the cells were either lysed for examining the total amount of internalized/biotinylated ALK (Tot-ALK at 0 min) or returned to 37ºC to resume normal endocytic trafficking (endocytosis/recycling/degradation). After 30 min incubation at 37 ºC, half of the cells were directly lysed (Tot-ALK at 30 min) while the other half were lysed after GSH treatment that removed any biotin from protein recycled back to the cell surface (IC-ALK at 30 min). The biotinylated ALK was pulled down using Streptavidin beads and blotted with HA antibody.

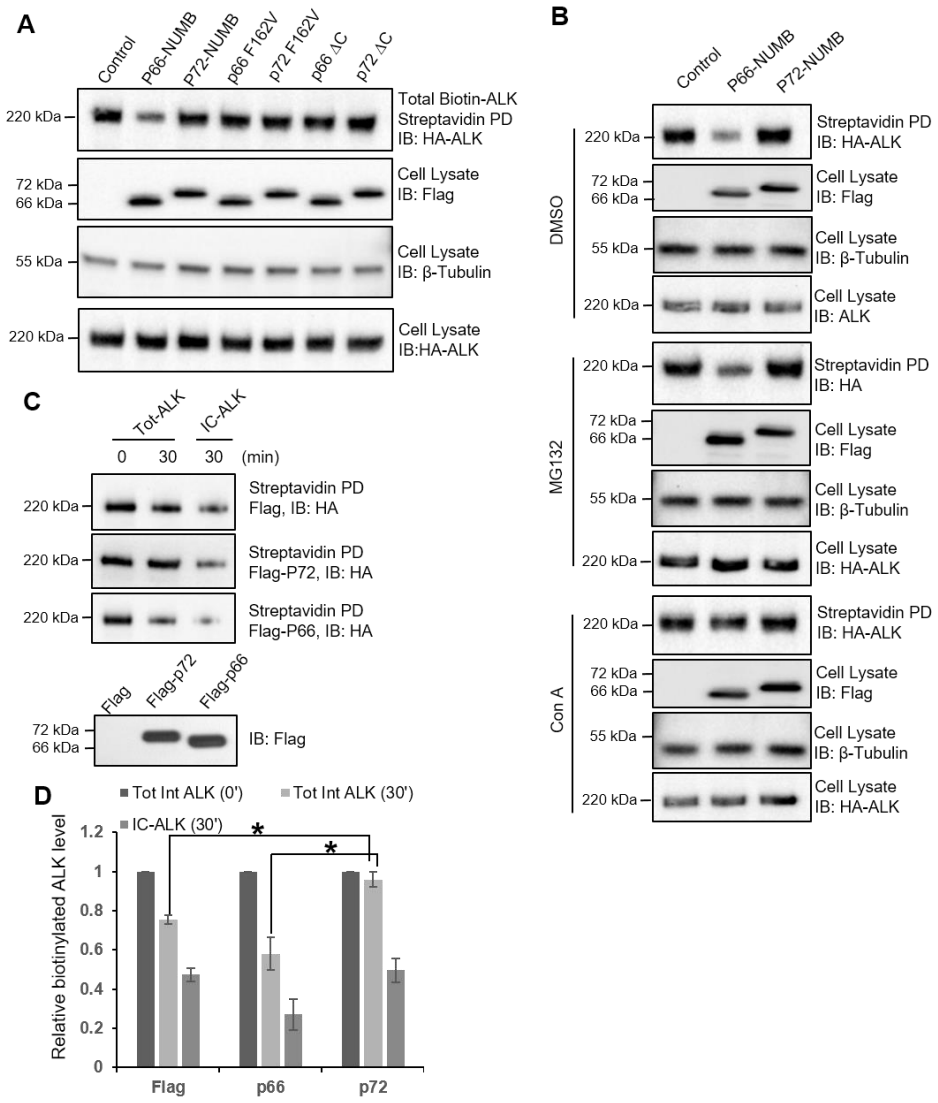


Figure.3.4 NUMB isoforms play distinct roles in ALK post-endocytic trafficking

(A) p66-NUMB, but not p72-NUMB, promoted post-endocytic ALK degradation. After surface biotinylation, the total biotinylated ALK level was determined in cells transfected with flag-tagged p66-NUMB and p72-NUMB isoforms, their F162V PTB binding-deficient mutants and C-terminal deletion forms lacking the endocytotic motif. Only p66-NUMB was found to decrease biotinylated ALK levels indicating degradation following endocytosis; (B) p66-NUMB promoted the degradation of endocytic ALK via the lysosomal pathway. The lysosome inhibitor concanamycin A attenuated p66-NUMB promoted biotinylated-ALK degradation, while DMSO and proteasome inhibitor MG132 had no obvious effect; (C-D) p72-NUMB, but not p66-NUMB, promoted post-endocytic ALK recycling. Surface biotinylation, internalization and recycling assays were carried out in cells transfected with either p66-NUMB, p72-NUMB or flag tag only. The p72-NUMB mainly maintained the level of biotinylated ALK (Tot-ALK 30 vs. Tot-ALK 0) and quantitatively recycled approximate 46% biotinylated ALK to the membrane (Tot-ALK 30 minus IC-ALK 30) in our test. Consistently, p66-NUMB expression resulted in more biotinylated ALK degradation. $P<0.05$, $n=3$.

As shown in Fig.3.4C, the Tot-ALK level was significantly decreased in cells expressing the p66 isoform after 30 min of endocytosis recovery at 37ºC which was consistent with the results presented in Fig.3.4A. However, only a minor change was observed in cells expressing the p72 isoform. A relative quantification of western blotting better illustrates the difference between the isoforms (Fig.3.4D). At the 30 min time point without the second GSH treatment, the biotinylated ALK level was ~96% of the initial level with p72 isoform expression while only ~58% with p66 isoform expression. Upon the second GSH treatment, the biotinylated ALK levels were ~50% and ~27% in cells expressing p72 or p66 respectively. This implies that p72 expression resulted ~46% (96%-50%) ALK recycling and p66 expression resulted ~31% (58%-27%) ALK recycling during our treatment.

To understand the mechanism underpinning the different functions of the two NUMB isoforms in ALK degradation and recycling, we examined their co-localizations with Rab4, Rab7, and Rab11 which were expressed as GFP-fused proteins in HEK293/ALK cells together with Flag-p66-NUMB or Flag-p72-NUMB. There was a significant increase in co-localization between ALK and the late endosome marker GFP-Rab7 (Chavrier et al., 1990) in HEK293/ALK cells co-expressing the p66, but not the p72 isoform of NUMB or the vector control (Fig. 3.5A). This suggests that p66-NUMB may facilitate the degradation of ALK by routing it to the late endosome and lysosome. In contrast, expression of p72-NUMB, but not p66-NUMB, promoted the co-localization of ALK with the early endosome marker Rab4 (Van Der Sluijs et al., 1991) (Fig. 3.5B). Because the recycling process is dependent on both Rab4 and Rab11 and/or Rab14 (Pfeffer, 2013), we also examined the co-localization of ALK with Rab11 in cells expressing either NUMB isoform. Intriguingly, neither NUMB isoform was able to significantly alter the co-localization of ALK and Rab11 (Fig. 3.5 C). These observations were further verified by a quantitative analysis of Pearson's co-efficiency from multiple cells/images respectively (Fig.3.5D), which exhibited significant differences in ALK-Rab4/7 co-localizations with the expression of p66 or p72 isoforms.

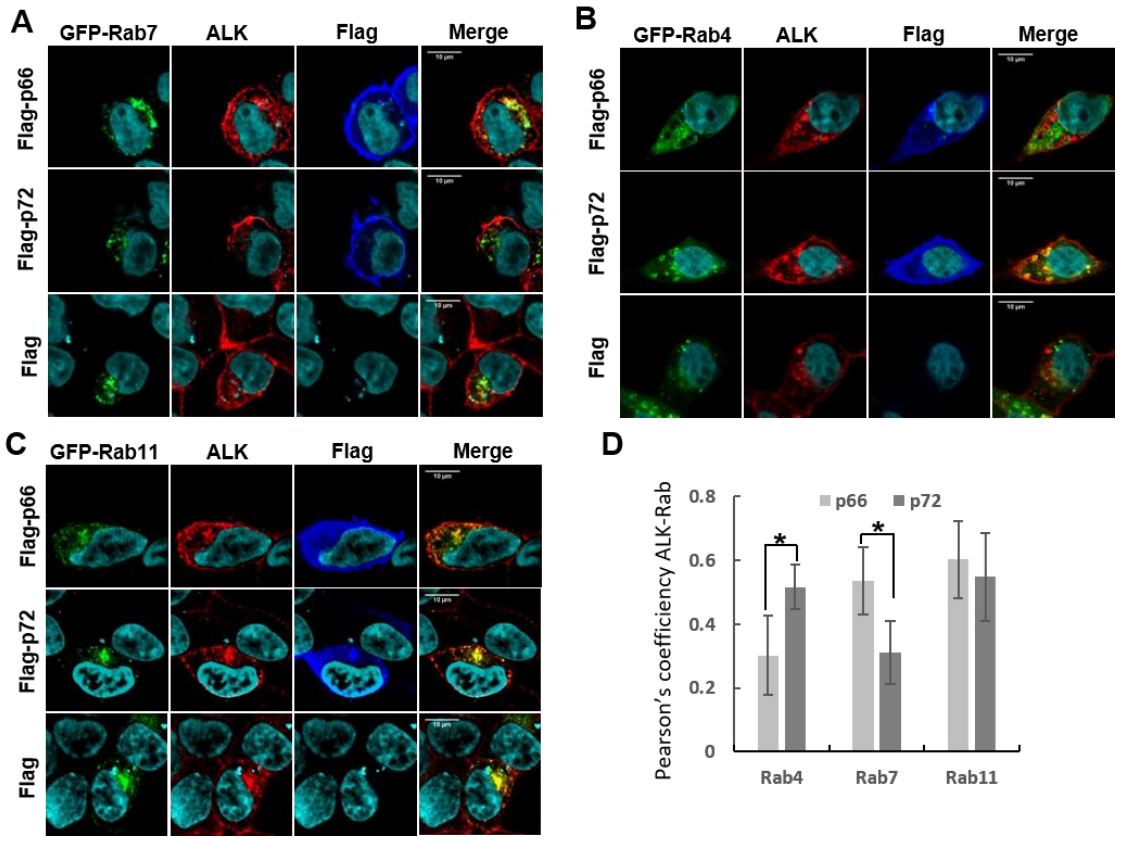


Figure.3.5 NUMB isoforms differentiate the co-localization between ALK and Rabs

HEK293/ALK cells were co-transfected with different combinations of flag labelled NUMB isoforms and GFP-fused Rab4/7/11 endocytic markers as presented in (A-C). Co-immunostaining was then carried out using anti-HA (ALK, red), anti-GFP (Rab, green) and anti-Flag (Numb, blue). Individual cells exhibiting all three fluorescent markers were picked out for further analysis. In the cytoplasm but not the cell membrane, p66-NUMB expression promoted the co-localization of ALK and the late endosome marker Rab7 (A), while the p72-NUMB expression promoted co-localization between ALK and early endosome marker Rab4 (B). Another early endosome marker, Rab11, generally co-localized to ALK in all conditions (C). These observations were further confirmed by the quantification of Pearson's co-efficiency (D). Rab4: $p<0.01$, $n=8$; Rab7: $p<0.01$, $n=5$; Rab11: $p<0.01$, $n=5$.

These data demonstrate that NUMB plays an important role in the sorting of endocytic ALK. p72-NUMB facilitates the fast-recycling of ALK via Rab4-containing early endosome, while p66-NUMB promotes the lysosome-dependent degradation of ALK via the Rab7-containing late endosome. NUMB regulates ALK activity via mediating ALK endocytosis

3.4.4 NUMB regulates ALK activity via mediating ALK endocytosis

ALK signals through the MAPK pathway to regulate cell proliferation (Hallberg and Palmer, 2013). To interrogate the role of two NUMB isoforms in ALK activity, we examined Erk phosphorylation in HEK293/ALK cells expressing either wildtype or ALK-binding deficient p66/p72. We used an antibody for the extracellular region of ALK, mAb46, to specifically stimulate the activation of ALK and MAP kinase pathway (Erk phosphorylation) under serum-free conditions (Moog-Lutz et al., 2005) . As shown in Fig.3.6A, Erk phosphorylation was promoted by mAb46 treatment in HEK293/ALK cells. mAb46 is the antibody specifically targeting ALK and can mimic ALK ligand to activate ALK. Co-expression of p66-NUMB significantly attenuated the Erk phosphorylation while its ALK-binding deficient form p66-F162V or endocytosis deficient form p66-ΔC had little effects (Fig.3.6A-B).

To further validate that the isoform-specific function of NUMB on ALK activation is dependent on their roles in ALK post-endocytic trafficking, we used the dominant-negative forms of Rab4/7/11 (GFP-Rab4/7/11DN) to specifically perturb endocytosis routes in HEK293/ALK co-expressed with Flag-p66-NUMB or Flag-p72-NUMB or Flag only. Because GFP-Rab7 DN has been shown to block endocytosis in the late endosome (Bucci et al., 2000), it is thought to prevent the internalized protein from lysosomal degradation. Indeed, we observed that the p66-NUMB mediated Erk phosphorylation level decrease (Fig.3.6C) was eliminated with GFP-Rab7DN expression (Fig.3.6D). This result suggested that the inhibitory effect of p66-NUMB on ALK activation is dependent on its ability to sort ALK to the late endosome.

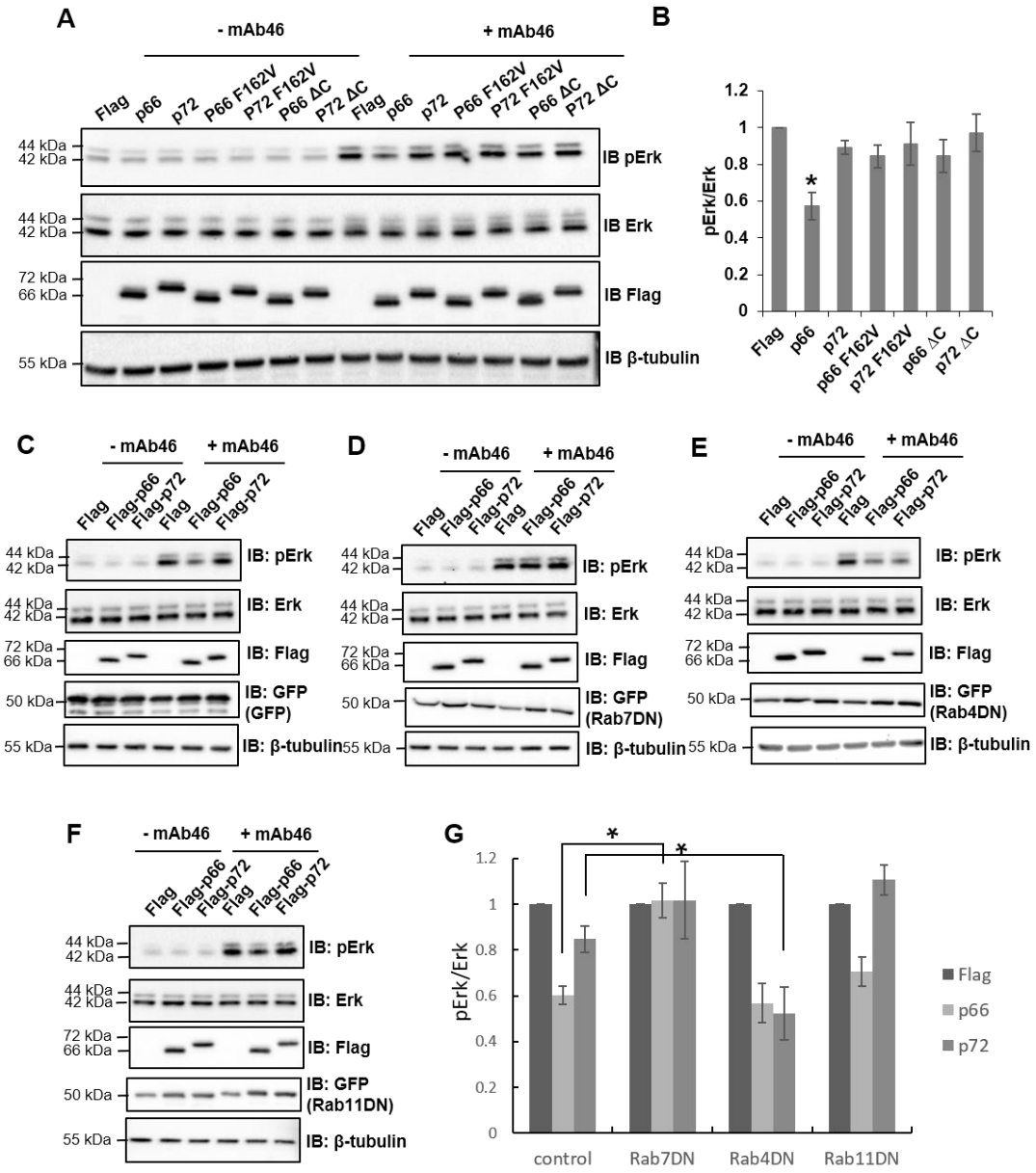


Figure.3.6 p66-NUMB inhibits ALK-mediated MAPK signaling

(A-B) Erk phosphorylation was attenuated by p66-NUMB expression in ALK activated HER293/ALK cells. ALK was selectively activated by mAb46 antibody treatment in serum starved cells. Only p66-NUMB, but not its ALK binding-deficient F162V mutant or C-terminal deletion form, decreased Erk phosphorylation levels. For the pErk/Erk ratio quantification in western blotting, $p<0.01$, n=3. (C-G) Interrupting endocytosis at different stages exhibited differential effects for p66-NUMB or p72-NUMB in ALK-pErk signal transduction. Overexpression of Rab7 dominant negative mutant blocked the endocytosis at late endosome, which eliminated p66-NUMB mediated ALK activity (Erk phosphorylation) decrease (D). Overexpression of the Rab4 dominant negative mutant blocked the recycling, which prevented p72-NUMB from maintaining ALK activity by recycling (E). Overexpression of empty vector (C) or Rab11 dominant negative mutant (F) had little effect. The changes in pErk/Erk ratio were significant, n=3, $p<0.01$ (G).

In the parallel tests with GFP-Rab4DN, we also observed that the block of fast recycling by GFP-Rab4DN (Yudowski et al., 2009) led to ~40% reduction in Erk phosphorylation in p72-NUMB expressing cells but had little effect in p66-NUMB expressing cells, demonstrating that p72-NUMB mediated ALK fast recycling promotes Erk phosphorylation (Fig.3.6E-F). In contrast, the block of slow recycling by GFP-Rab11DN had no significant effect on Erk activation in cells co-expressing either p66-Numb or p72-Numb (Fig.3.6G-H).

Because ALK usually exists as a constitutively active form in tumors, we also evaluated the NUMB mediated ALK activity with ALK-F1174L mutant. ALK-WT and ALK-F1174L were transiently expressed in HEK293 cells. As shown in Fig.3.7A, both ALK forms stimulated the activation of MAPK pathway under normal cell culturing conditions with 10% serum, while the ALK-F1174L had significant higher Erk phosphorylation level as expected. This Erk phosphorylation was only attenuated by co-expression p66-NUMB but not p72-NUMB. Consistently, in the cell proliferation assay and soft-agar assay with ALK-F1174L transfected HEK293 cells, only co-expression of p66-NUMB exhibited a moderate inhibitory effect (Fig.3.7B-C). In conclusion, p66-NUMB and p72-NUMB act antagonistically in regulating ALK and MAPK activities via routing ALK to different endocytic pathways.

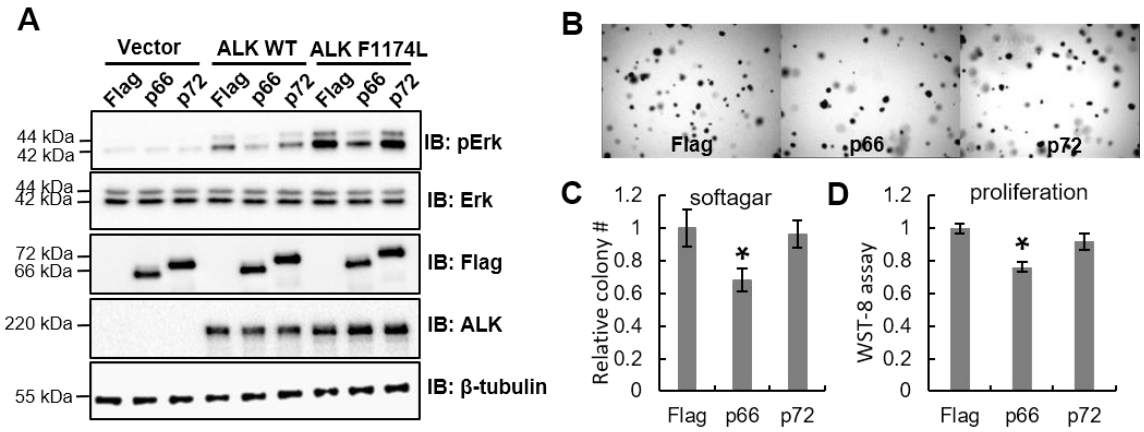


Figure.3.7 p66-NUMB inhibits ALK-dependent cancer cell growth

(A) p66-NUMB inhibited the activation of both ALK and ALK oncogenic mutant F1174L. The empty vectors, ALK wild type, ALK constitutively active mutant F1174L and NUMB isoforms were transiently expressed in HEK293 cells. After cell starvation, ALK was specifically stimulated by antibody. Both ALK WT and ALK-F1174L were able to increase the phosphorylation of Erk, which was only attenuated by co-expression of p66-NUMB. (B-C) p66-NUMB inhibited the anchorage-independent cell growth of HEK293/ALK cells, n=3, $p<0.01$. Empty vector, p66-NUMB and p72-NUMB were transiently expressed in HEK293/ALK cells. 24 hours after transfection, the cells were used in a soft agar assay. (D) p66-NUMB inhibited the proliferation of HEK293/ALK cells. WST-8 assay was carried out 48 hours after transfection, n=3, $p<0.01$.

3.5 Discussion

The NUMB protein was the first described cell fate determinant which is conserved from fly to human. Further studies have since revealed the complex nature of NUMB functions that are involved in a wide range of fields. In tumorigenesis, NUMB is defined as a tumor suppressor for its role in binding and regulation of the NOTCH receptor and p53 (Colaluca et al., 2008; Couturier et al., 2012; McGill et al., 2009; McGill and McGlade, 2003). Interestingly, the NUMB-NOTCH and NUMB-p53 interactions take place in the cell membrane periphery and nucleus, and suppress tumorigenesis via distinct mechanisms, exhibiting the multifaceted functional pattern of NUMB. NUMB function mainly relies on interactions between its modular PTB domain and binding partners (NxxF/NxxY motifs). As presented in Chapter 2, the PTB domain mediates interactions between NUMB and a large number of proteins. On the other hand, the non-PTB region of NUMB seems to play a more significant role for the NUMB functional diversity. The PTB^L and PTB^S isoforms exhibit little difference other than distinct sub-cellular localizations, however, the PRR^L and PRR^S isoforms are well known for their different functions in proliferation and differentiation during development (Bani-Yaghoub et al., 2007; Dho et al., 1999; Dooley et al., 2003; Verdi et al., 1999a; Yoshida et al., 2003). Consistently, our findings demonstrate that both PTB^L and PTB^S isoforms bind to two ALK motifs that are required for the NUMB-ALK interaction. However, the PRR^L isoform p72 mediates ALK recycling and PRR^S isoform p66 mainly promotes ALK degradation, resulting in an increase in cell proliferation MAPK signaling with p72 isoform expression compared to that with p66 isoform expression. Based on microscopy data, it is clearly presented that ALK co-localizes to different post-endocytic destination makers with p66-NUMB or p72-NUMB expression. Our study first methodically elucidates the functional diversity between PRR^L and PRR^S isoforms in the regulation of ALK endocytosis and endosomal sorting (Fig.3.8). Previously, it was reported that NUMB is involved in both recycling and endosomal compartmentalization of NOTCH (McGill et al., 2009). Therefore, it is possible that the NUMB isoforms may act in a similar manner for all NUMB-binding receptors (including RTKs). Our finding is consisting with reported NUMB isoform-dependent oncogenic behaviors (Bechara et al., 2013; Misquitta-Ali et al., 2011; Westhoff et al., 2009), which

further raises the possibility the p72-NUMB may universally assist in sustaining receptor activity thus contributing to tumorigenesis. Considering receptors are constituting a major portion of the NUMB-PTB binding partners, as presented in Chapter2, our study also has potential value for the improvement of diagnosis and therapy design of receptor positive cancers.

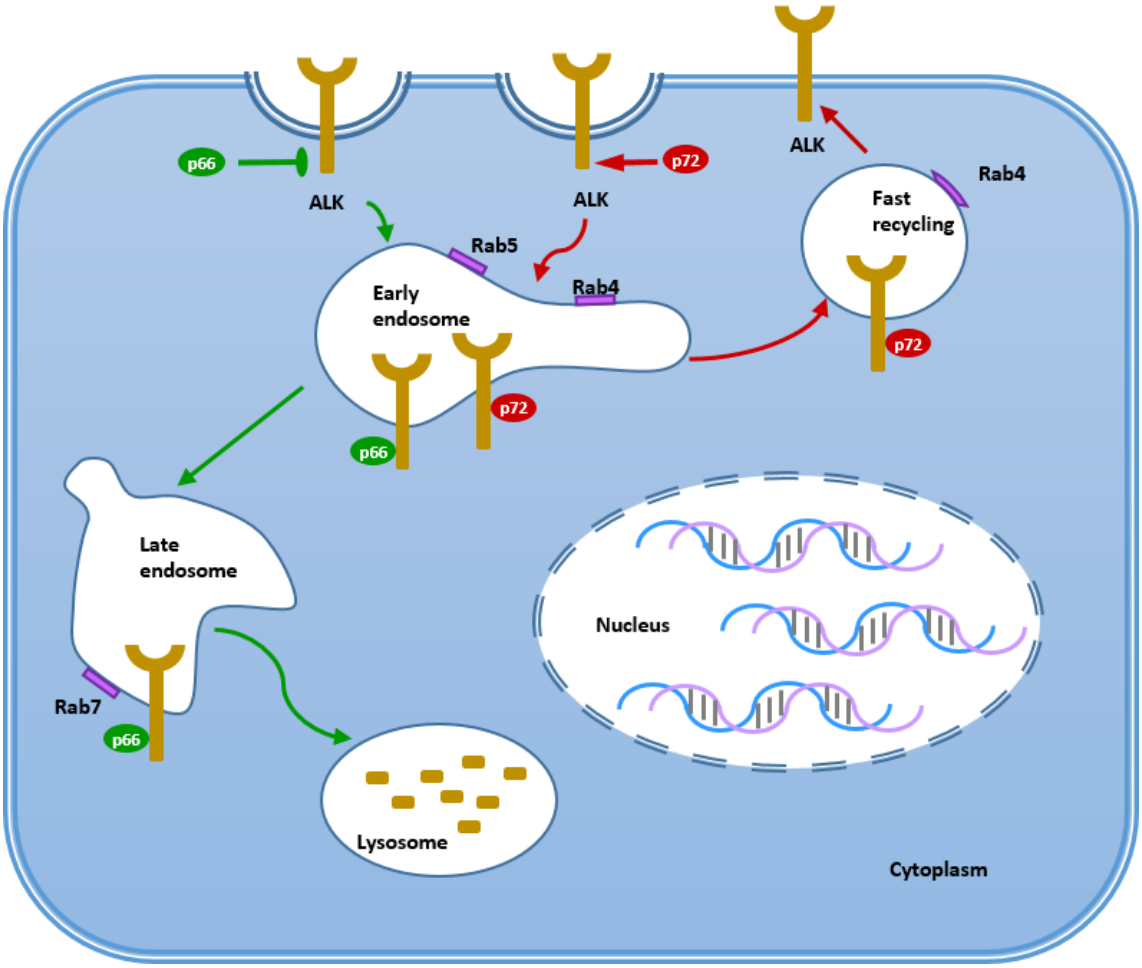


Figure.3.8 Overview of NUMB-regulated ALK endocytosis

The mechanisms by which NUMB isoforms regulate ALK endosomal sorting have not been well defined. Based on preliminary findings, we hypothesize that p72-NUMB and p66-NUMB may recruit different endosomal sorting factors following ALK internalization. It must be noted that p72-NUMB and p66-NUMB do not play absolutely opposing roles in the ALK endosomal sorting but more likely prefer the recycling or late

endosomal compartment respectively. A mass spectrometry based approach has been employed to quantitatively compare the interactome between p72-NUMB and p66-NUMB. Indeed, it was found that both isoforms generally bind to several known NUMB-interacting endocytic proteins, but p72-NUMB bound significantly more EPS15 and AP-2 (Krieger et al., 2013). These data confirm the possibility that NUMB isoforms may act differently during endocytosis and also reveals some additional effectors that may be involved in these functional differentiations. Overall, further mechanistic studies will be required to determine the contribution of NUMB-binding proteins in both internalization and post-endocytic trafficking events.

3.6 References

Bani-Yaghoub, M., Kubu, C. J., Cowling, R., Rochira, J., Nikopoulos, G. N., Bellum, S., and Verdi, J. M. (2007). A switch in numb isoforms is a critical step in cortical development. Developmental Dynamics *236*, 696-705.

Bavi, P., Jehan, Z., Bu, R., Prabhakaran, S., Al-Sanea, N., Al-Dayel, F., Al-Assiri, M., Al-Halouly, T., Sairafi, R., Uddin, S., and Al-Kuraya, K. S. (2013). ALK gene amplification is associated with poor prognosis in colorectal carcinoma. Br J Cancer *109*, 2735-2743.

Bechara, E. G., Sebestyen, E., Bernardis, I., Eyras, E., and Valcarcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Molecular cell *52*, 720-733.

Bilsland, J. G., Wheeldon, A., Mead, A., Znamenskiy, P., Almond, S., Waters, K. A., Thakur, M., Beaumont, V., Bonnert, T. P., Heavens, R., *et al.* (2008). Behavioral and neurochemical alterations in mice deficient in anaplastic lymphoma kinase suggest therapeutic potential for psychiatric indications. Neuropsychopharmacology *33*, 685-700.

Bork, P., and Margolis, B. (1995). A phosphotyrosine interaction domain. Cell *80*, 693-694.

Bucci, C., Thomsen, P., Nicoziani, P., McCarthy, J., and van Deurs, B. (2000). Rab7: a key to lysosome biogenesis. Molecular biology of the cell *11*, 467-480.

Chavrier, P., Parton, R. G., Hauri, H. P., Simons, K., and Zerial, M. (1990). Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. Cell *62*, 317-329.

Colaluca, I. N., Tosoni, D., Nuciforo, P., Senic-Matuglia, F., Galimberti, V., Viale, G., Pece, S., and Di Fiore, P. P. (2008). NUMB controls p53 tumour suppressor activity. Nature *451*, 76-80.

Couturier, L., Vodovar, N., and Schweisguth, F. (2012). Endocytosis by Numb breaks Notch symmetry at cytokinesis. Nat Cell Biol *14*, 131-139.

Crescenzo, R., and Inghirami, G. (2015). Anaplastic lymphoma kinase inhibitors. Curr Opin Pharmacol *23*, 39-44.

Dho, S. E., French, M. B., Woods, S. A., and McGlade, C. J. (1999). Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. J Biol Chem *274*, 33097-33104.

Dooley, C. M., James, J., McGlade, C. J., and Ahmad, I. (2003). Involvement of numb in vertebrate retinal development: Evidence for multiple roles of numb in neural differentiation and maturation. Journal of Neurobiology *54*, 313-325.

George, R. E., Sanda, T., Hanna, M., Frohling, S., Luther, W., 2nd, Zhang, J., Ahn, Y., Zhou, W., London, W. B., McGrady, P., *et al.* (2008). Activating mutations in ALK provide a therapeutic target in neuroblastoma. Nature *455*, 975-978.

Goh, L. K., and Sorkin, A. (2013). Endocytosis of receptor tyrosine kinases. Cold Spring Harb Perspect Biol *5*, a017459.

Gulino, A., Di Marcotullio, L., and Screpanti, I. (2010). The multiple functions of Numb. Exp Cell Res *316*, 900-906.

Hallberg, B., and Palmer, R. H. (2013). Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. Nature reviews Cancer *13*, 685-700.

Kaushansky, A., Allen, J. E., Gordus, A., Stiffler, M. A., Karp, E. S., Chang, B. H., and MacBeath, G. (2010). Quantifying protein-protein interactions in high throughput using protein domain microarrays. Nat Protoc *5*, 773-790.

Krieger, J. R., Taylor, P., Gajadhar, A. S., Guha, A., Moran, M. F., and McGlade, C. J. (2013). Identification and selected reaction monitoring (SRM) quantification of endocytosis factors associated with Numb. Mol Cell Proteomics *12*, 499-514.

Li, S. C., Zwahlen, C., Vincent, S. J., McGlade, C. J., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (1998). Structure of a Numb PTB domain-peptide complex suggests a basis for diverse binding specificity. Nat Struct Biol *5*, 1075-1083.

McGill, M. A., Dho, S. E., Weinmaster, G., and McGlade, C. J. (2009). Numb regulates post-endocytic trafficking and degradation of Notch1. J Biol Chem *284*, 26427-26438.

McGill, M. A., and McGlade, C. J. (2003). Mammalian numb proteins promote Notch1 receptor ubiquitination and degradation of the Notch1 intracellular domain. J Biol Chem *278*, 23196-23203.

Misquitta-Ali, C. M., Cheng, E., O'Hanlon, D., Liu, N., McGlade, C. J., Tsao, M. S., and Blencowe, B. J. (2011). Global profiling and molecular characterization of alternative splicing events misregulated in lung cancer. Mol Cell Biol *31*, 138-150.

Moog-Lutz, C., Degoutin, J., Gouzi, J. Y., Frobert, Y., Brunet-de Carvalho, N., Bureau, J., Creminon, C., and Vigny, M. (2005). Activation and inhibition of anaplastic lymphoma kinase receptor tyrosine kinase by monoclonal antibodies and absence of agonist activity of pleiotrophin. J Biol Chem *280*, 26039-26048.

Morris, S. W., Kirstein, M. N., Valentine, M. B., Dittmer, K. G., Shapiro, D. N., Saltman, D. L., and Look, A. T. (1994). Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science *263*, 1281-1284.

Pece, S., Confalonieri, S., P, R. R., and Di Fiore, P. P. (2011). NUMB-ing down cancer by more than just a NOTCH. Biochim Biophys Acta *1815*, 26-43.

Pfeffer, S. R. (2013). Rab GTPase regulation of membrane identity. Curr Opin Cell Biol *25*, 414-419.

Salcini, A. E., Confalonieri, S., Doria, M., Santolini, E., Tassi, E., Minenkova, O., Cesareni, G., Pelicci, P. G., and Di Fiore, P. P. (1997). Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. Genes Dev *11*, 2239-2249.

Salido, M., Pijuan, L., Martinez-Aviles, L., Galvan, A. B., Canadas, I., Rovira, A., Zanui, M., Martinez, A., Longaron, R., Sole, F., *et al.* (2011). Increased ALK gene copy number and amplification are frequent in non-small cell lung cancer. J Thorac Oncol *6*, 21-27.

Sasaki, T., Rodig, S. J., Chirieac, L. R., and Janne, P. A. (2010). The biology and treatment of EML4-ALK non-small cell lung cancer. Eur J Cancer *46*, 1773-1780.

Shiota, M., Fujimoto, J., Semba, T., Satoh, H., Yamamoto, T., and Mori, S. (1994). Hyperphosphorylation of a novel 80 kDa protein-tyrosine kinase similar to Ltk in a human Ki-1 lymphoma cell line, AMS3. Oncogene *9*, 1567-1574.

Smith, C. A., Dho, S. E., Donaldson, J., Tepass, U., and McGlade, C. J. (2004). The cell fate determinant numb interacts with EHD/Rme-1 family proteins and has a role in endocytic recycling. Molecular biology of the cell *15*, 3698-3708.

Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., *et al.* (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature *448*, 561-566.

Tumilowicz, J. J., Nichols, W. W., Cholon, J. J., and Greene, A. E. (1970). Definition of a continuous human cell line derived from neuroblastoma. Cancer Res *30*, 2110-2118.

Uemura, T., Shepherd, S., Ackerman, L., Jan, L. Y., and Jan, Y. N. (1989). numb, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos. Cell *58*, 349-360.

Van Der Sluijs, P., Hull, M., Zahraoui, A., Tavitian, A., Goud, B., and Mellman, I. (1991). The small GTP-binding protein rab4 is associated with early endosomes. Proceedings of the National Academy of Sciences of the United States of America *88*, 6313-6317.

Verdi, J., Schmandt, R., Bashirullah, A., Jacob, S., Salvino, R., Craig, C. G., Lipshitz, H. D., and McGlade, C. J. (1996). Mammalian NUMB is an evolutionarily conserved signaling adapter protein that specifies cell fate. Curr Biol *6*, 1134-1145.

Verdi, J. M., Bashirullah, A., Goldhawk, D. E., Kubu, C. J., Jamali, M., Meakin, S. O., and Lipshitz, H. D. (1999a). Distinct human NUMB isoforms regulate differentiation vs. proliferation in the neuronal lineage. Proceedings of the National Academy of Sciences of the United States of America *96*, 10472-10476.

Wang, M., Zhou, C., Sun, Q., Cai, R., Li, Y., Wang, D., and Gong, L. (2013). ALK amplification and protein expression predict inferior prognosis in neuroblastomas. Exp Mol Pathol *95*, 124-130.

Wang, Z., Sandiford, S., Wu, C., and Li, S. S. (2009). Numb regulates cell-cell adhesion and polarity in response to tyrosine kinase signalling. The EMBO journal *28*, 2360-2373.

Westhoff, B., Colaluca, I. N., D'Ario, G., Donzelli, M., Tosoni, D., Volorio, S., Pelosi, G., Spaggiari, L., Mazzarol, G., Viale, G., *et al.* (2009). Alterations of the Notch pathway in lung cancer. Proc Natl Acad Sci U S A *106*, 22293-22298.

Yoshida, T., Tokunaga, A., Nakao, K., and Okano, H. (2003). Distinct expression patterns of splicing isoforms of mNumb in the endocrine lineage of developing pancreas. Differentiation *71*, 486-495.

Yudowski, G. A., Puthenveedu, M. A., Henry, A. G., and von Zastrow, M. (2009). Cargo-mediated regulation of a rapid Rab4-dependent recycling pathway. Molecular biology of the cell *20*, 2774-2784.

Zwahlen, C., Li, S. C., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (2000). Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. The EMBO journal *19*, 1505-1515.

Chapter 4

4 ALK ectodomain shedding

4.1 Abstract

The human ALK gene encodes a 180-kDa protein consisting of 1620 amino acids, which corresponds to a 220-kDa band on SDS-PAGE. In addition, the 140-kDa ALK band is frequently observed upon immunoblotting, suggests that ALK may undergo ectodomain shedding, a process employed by many different receptors to generate intracellular truncated fragments. Due to diversified subcellular localizations and binding partners, these truncated fragments exhibit a functional pattern distinct from that of the full-length receptor on the cell surface. To further investigate the potential function of the truncated ALK form, we identified the ALK cleavage regions by analyzing the ectodomain sequence using mass spectrometry. In addition, we found that the 140-kDa fragment is highly phosphorylated even in the absence of ligand-stimulated activation, suggesting ectodomain shedding may contribute to ALK oncogenicity by generating this cytoplasmic fragment.

4.2 Introduction

Ectodomain shedding is a proteolytic cleavage of membrane proteins that results in the release of the ectodomain. This mechanism modulates a wide variety of cellular and physiological processes, including growth factor signaling, cell adhesion and inflammation (Hayashida et al., 2010; Weber and Saftig, 2012). It is estimated that cells selectively shed only 2% of membrane proteins, indicating the existence of a regulatory system for this process (Hayashida et al., 2010). Interestingly, ectodomain shedding is common in RTKs and has been reported in members of almost all RTK subfamilies (Chen and Hung, 2015; Miller et al., 2013).

Ectodomain shedding is typically carried out by the matrix metalloproteinase (MMP) family or the a disintegrin and metalloproteinase (ADAM) family. MMPs are secreted proteases capable of degrading variety of proteins in the extracellular matrix (ECM) and can also cleave certain membrane proteins. However, ectodomain shedding by MMPs is not very tightly regulated, and mainly depends on MMP abundance in the ECM (Murphy, 2008). ADAM proteases are transmembrane proteases and termed as “sheddases” that specifically cleave ectodomains at the cell surface (Sternlicht and Werb, 2001). Unlike that of MMPs, ADAM-mediated ectodomain shedding is found to be regulated by different types of stimulation, including undefined serum factors, peptide growth factors, calcium concentration and some forms of stress (Seals and Courtneidge, 2003). Although MMPs and ADAMs have distinct characteristics, both families are physiologically functional in tumorigenesis (Duffy et al., 2009; Lu et al., 2011), partially due to their involvement in the processing of receptors and receptor ligands. For instance, ectodomain shedding of either ErbB or its ligand precursors plays a crucial role in the regulation of ErbB oncogenicity. ADAMs mediate the release of the EGF ligand from its precursor HB-EGF membrane protein, which is required for the activation of ErbB RTKs (Blobel, 2005; Revillion et al., 2008). ErbB itself also undergoes ectodomain shedding, generating a soluble ErbB fragment in the ECM. This free ErbB fragment has been defined as a potential biomarker for fast diagnosis of breast and ovarian cancers (Lafky et al., 2008; Visco et al., 2000). Besides ectodomain shedding, some receptors undergo further proteolytic cleavage. As introduced in Chapter 1, the maturation of the NOTCH receptor requires intracellular

cleavage by γ-secretase following ectodomain shedding. An identical process is also observed in ErbB2 and some other RTKs (Chen and Hung, 2015). In the cytoplasm, free truncated ErbBs may function similarly to endosomal ErbB by stimulating signal transduction. In addition, the ErbB2 intracellular region contains a typical nuclear localization sequence (NLS) that mediates ErbB2 nuclear translocation, where it functions as a transcription co-activator (Giri et al., 2005).

ALK also undergoes ectodomain shedding. The human ALK gene encodes a 180-kDa protein that consist of 1620 amino acids. Upon posttranslational modification, in particular N-linked glycosylation, the mature ALK protein migrates as a 220-kDa band on SDS-PAGE (Mazot et al., 2011). ALK also exists as a 140-kDa form, due to proteinase cleavage at the extracellular region, but the cleavage site is still undetermined (Moog-Lutz et al., 2005). Here we present preliminary data for the investigation of ALK cleavage. We identified the ALK cleavage region and found the 140-kDa is highly phosphorylated even in the absence of ligand-stimulated activation.

4.3 Material and methods

4.3.1 Cell Culture and treatment

HEK-293/ALK is a HEK cell line stably expressing ALK. This cell line was cultured in DMEM media supplemented with 10% FBS, 2 mg/ml glutamine, 100 units of penicillin-streptomycin and 2 µg/ml puromycin. HEK 293 cells were cultured in the same medium except without puromycin. IMR-32 cells were cultured in RPMI 1640 media supplemented with 10% FBS, 2 mg/ml glutamine, 100 units of penicillin-streptomycin. 2 µM BB-94 was added into the HEK293-ALK cell culture media. After overnight treatment, the cell lysate was tested for ALK by western blot. For PMA, APMA and GM6001 treatments, the HEK293-ALK cells were first washed with 1×PBS buffer. Next, the chemicals were diluted to the designated concentrations in DMEM to treat the cells. Under these conditions, both medium and cell lysates were collected then western blots were performed against ALK. Anti-N-terminal ALK antibody was used to detect ALK in the DMEM media. Anti-HA (1:1000) or Anti-C-terminal ALK (1:500) antibody was used to detect ALK in the cell lysate. The HEK293-ALK cells were treated with 1 mM, 0.5 mM and 0.25 mM APMA for 30 minutes to determine the optimal concentration for ALK cleavage. Next, the HEK293-ALK cells were incubated for 30 minutes either with DMEM alone, 0.5 mM APMA, 0.5 mM APMA in combination with different GM6001 concentrations (10 µM & 50 µM), or PMA at different concentrations (10 µM & 100 µM). Both media and cell lysates were collected for western blot analysis.

4.3.2 PCR mediated ALK truncations

The full length human ALK with a C-terminal HA tag was expressed in a pCBC vector (This plasmid was a generous gift from Dr. Marc Vigny). Truncated ALKs: ALK 509-552 Del, ALK 552-612 Del, ALK 21-400 Del, ALK 21-700 Del and ALK 21-1031 Del were generated by PCR-mediated plasmid DNA deletion (Hansson et al., 2008). Primers were designed based on the same principles as site-directed mutagenesis: 16 to 20 nucleic acids upstream of the deleted region followed by 16 to 20 nucleic acids downstream. The primers for PCR amplification of the listed ALK truncations are as follows:

ALK 509-552 Del F: CACGAATGAGCCAGGACATGGCATCCTTTAGGGTCCT

ALK509-552 Del R: AGGACCCTAAAGGATGCCATGTCCTGGCTCATTCGTG

ALK 552-612 Del F: TTGTCCCCACCATGCGACTCGGAGCTCACATGGAGA

ALK 552-612 Del R: TCTCCATGTGAGCTCCGAGTCGCATGGTGGGGACAA

ALK 21-400 Del F: TGTATTCCAGGGCCACTCGGCCCACAGCTGCCGTGGAA

ALK 21-400 Del R: TTCCACGGCAGCTGTGGGCCGAGTGGCCCTGGAATACA

ALK 21-700 Del F: TCTGGTAGGCGTTGTTGCAGCCCACAGCTGCCGTGGAA

ALK 21-700 Del R: TTCCACGGCAGCTGTGGGCTGCAACAACGCCTACCAGA

ALK 21-1030 Del F: GGATCAGCGAGAGTGGCAGGCCCACAGCTGCCGTGGAA

ALK 21-1030 Del R: TTCCACGGCAGCTGTGGGCCTGCCACTCTCGCTGATCC

The reaction was carried out in a mixture containing: 22.5 µL of H_2O, 2 µL of Pfu-Turbo DNA Polymerase (Agilent, Cat#600250), 7.5 µL of Pfu-Turbo PCR buffer, 10 µL of 2 mM dNTP, 2 µL of 125 ng/µL primers, 1.5 µL of DNA template (pCBC-ALK plasmid), and 2.5 µL of DMSO (50 µL total reaction volume), with 30 µL mineral oil added on top. The PCR reactions were performed with the following conditions: denaturation at 94°C for 2 minutes; followed by 18 cycles of 94°C denaturation for 50 seconds, 58°C annealing for 1 minute and 68°C elongation for 30 minutes; ending with 30 minutes of additional elongation. Next, 1 µL of Dpn1 was added to the PCR reaction products which were incubated at 37°C overnight. Then, 8 µL of the Dpn1-digested PCR products were used to transform DH5α *E. coli* competent cells. The plasmid DNA from these bacterial cultures were purified and sent for sequencing to verify successful ALK truncation.

4.3.3 Immunoprecipitation and western blotting

The cultured cells were lysed in cold mammalian cell lysis buffer (1% Triton X-100, 50 mM Tris-pH7.2, 150 mM NaCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM EGTA, 0.5 mM DTT, 10 mM NaF) containing complete protease inhibitors (Roche). Cell debris was removed by centrifugation, and 500 µg of supernatant was incubated for four hours in the presence of 1µg antibody and 30 µl of 50% slurry Protein G-Sepharose beads (Roche) at

4°C. The beads were subsequently washed three times with lysis buffer and boiled with SDS-loading dye. For western blotting, the immuno-precipitates were resolved by SDS-PAGE. All proteins were then transferred to a PVDF membrane by semi-dry transferring method (Bio-Rad), detected by immunoblotting with appropriate antibodies, and visualized by Enhanced Chemiluminescence (ECL). For mass spectrometry analysis, the precipitated protein was dissolved in 8 M urea for in-solution digestion.

4.3.4 Protein digestion and mass spectrometry analysis

The precipitated protein was first dissolved in 8 M urea prepared in MS-grade water. After reduction by dithiothreitol (DTT) and alkylation by iodoacetamide (IAA), the protein sample was digested by trypsin, chymotrypsin and Glu-C according to the manufacturer's protocol. Digested products were analyzed by Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometry equipped with an EASY-nLC liquid chromatography peptide separation system. All detected peptides were profiled and then aligned with their ALK protein sequence.

4.4 Results

4.4.1 Identification of ALK-140-kDa C-terminal fragment

The expression of ALK was determined in two cell lines: the HEK293/ALK stable cells that overexpress wildtype ALK with a C-terminal HA tag, and IMR-32 neuroblastoma cells that are known for endogenous ALK expression. As shown in Fig.4.1A, two bands were detected in both cell lines, and the 140-kDa band contains the ALK C-terminal region. Considering the molecular weight of this ALK fragment, we estimated that the cleavage site may be located within amino acid 509 to 612 (aa509 ~ aa612) region. Therefore, a number of truncated ALK forms were designed (Fig.4.1B) and transiently expressed with a C terminal HA tag in HEK293 cells. As shown in Fig.4.1C, both ALK aa509~aa552 and ALK aa552~aa612 deletion forms display the 140-kDa band by western blot analysis, suggesting that the cleavage site is located outside of aa509-aa612.

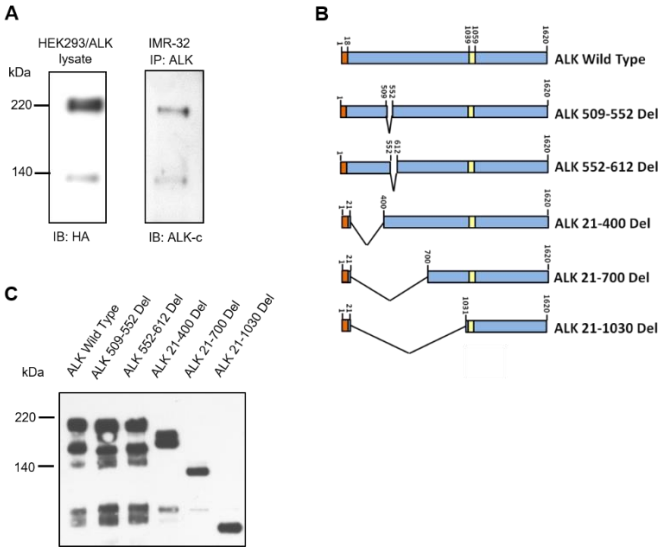


Figure.4.1 Identification of ALK-140-kDa fragment

(A) ALK 220-kDa and 140-kDa bands were present in both HEK293/ALK-HA stable cells and IMR-32 neuroblastoma cells. Considering the high ALK expression level in HEK293/ALK cells, the cell lysate was directly blotted with anti-HA antibody. For IMR-32 cells, endogenous ALK was first immunoprecipitated with ALK antibody and then blotted with ALK C-terminal antibody; (B-C) the cleavage site was predicted to be within aa509 to aa612 region. Several ALK truncated forms were designed (B) and transiently expressed in HEK293 cells, however both aa509~a552 and aa552-aa612 ALK deletion forms generated the 140-kDa fragment *in vivo*.

Figure.4.2 MS analysis of ALK ectodomain

An overview of ALK ectodomain peptides identified in MS. All detected peptides, which were generated from digestion by three different proteases, are aligned with the ALK sequence.

4.4.2 Screening for ALK cleavage site by MS

We used mass spectrometry to further identify the cleavage site by screening for peptides throughout the ALK ectodomain. In a serum-free environment, cultured HEK293/ALK cells were stimulated with 0.5 mM MMP/ADAM activator APMA for 30 minutes to

release the ALK ectodomain to the medium, which was then immuno-precipitated with an ALK N-terminal antibody. The product was digested in-solution by three different enzymes: chymotrypsin, trypsin and Glu-C, then all resulting peptides were profiled by mass spectrometry. The detected peptides were aligned with ALK protein sequence as shown in Fig.4.2 and Suppl.S7-9; the blue bars represent peptides with high signal intensities. From this data, it indicates that the cleavage site is located after, but close to amino acid 656.

4.4.3 MMP and ADAM mediate ALK ectodomain shedding

Receptor ectodomain shedding is typically carried out by metalloproteinase MMPs and ADAMs. To confirm that ALK cleavage is indeed mediated by this mechanism, we compared the effects of MMP/ADAM inhibitor (BB-94) and activator (APMA) in generating the 140-kDa ALK fragment in HEK293/ALK stable cells. We found the relative levels of 140-kDa ALK were significantly reduced with BB-94 treatment (Fig.4.3A-B) and increased with APMA treatment (Fig.4.3C). We also tested another broad MMP/ADAM inhibitor GM6001. At two different concentrations, 10 μM and 50 μM, GM6001 completely inhibited the effect of APMA and abolished ALK ectodomain shedding since neither the ALK ectodomain nor 140-kDa fragment were detected in the culture medium or cell lysate (Fig.4.3D, lane 1-4). In addition, we confirmed the ectodomain shedding of ALK was promoted by the activation of Protein Kinase C (PKC) with the treatment of PKC activator phorbol-12-myristate-13-acetate (PMA) (Fig.4.3D, lane 1, 5-6). This is consistent with previous reports that PKC is a major activator of ectodomain shedding (Hayashida et al., 2010).

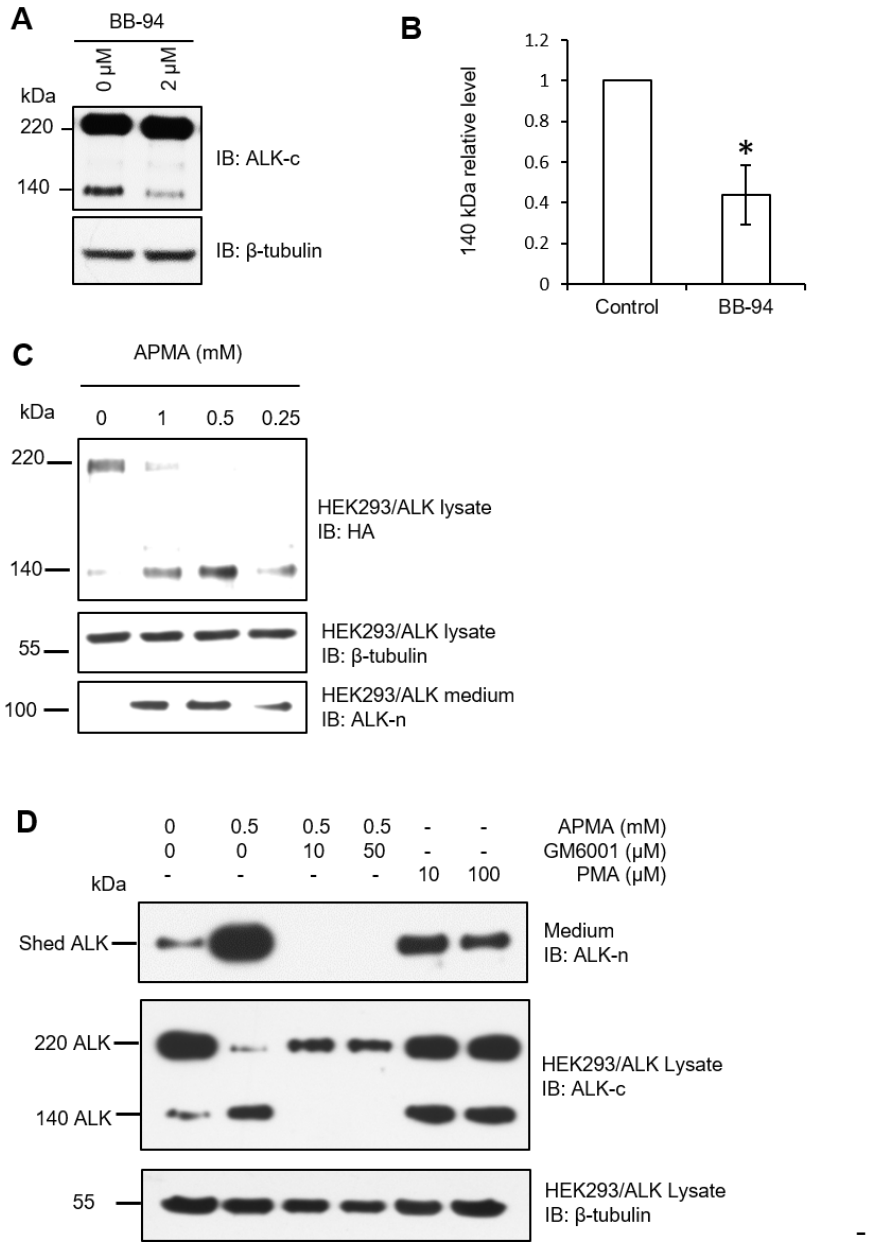


Figure.4.3 MMPs mediate ALK ectodomain shedding

(A) MMP/ADAM inhibitor BB-94 inhibited the generation of the 140-kDa ALK fragment; (B) The western blotting quantification of relative 140-kDa ALK fragment level with BB-94 treatment. n=3, $p<0.05$; (C) MMP/ADAM activator APMA promoted the generation of the 140-kDa ALK fragment; (D) MMP/ADAM broad inhibitor GM6001 completely blocked ALK ectodomain shedding (lane1-4). PKC activator PMA stimulated ALK ectodomain shedding (lane1, 5-6)

4.4.4 140-kDa ALK fragment is highly phosphorylated

The intracellular RTK fragments usually exhibit higher kinase activity or phosphorylation level compared to their intact forms in the cell membrane. We compared the tyrosine phosphorylation levels of 140-kDa and full length ALK isolated from the same cell lysate. As predicted, the 140-kDa ALK exhibited a much higher phosphorylation level compared to the full length ALK during APMA activation (Fig.4.4 lane 1).

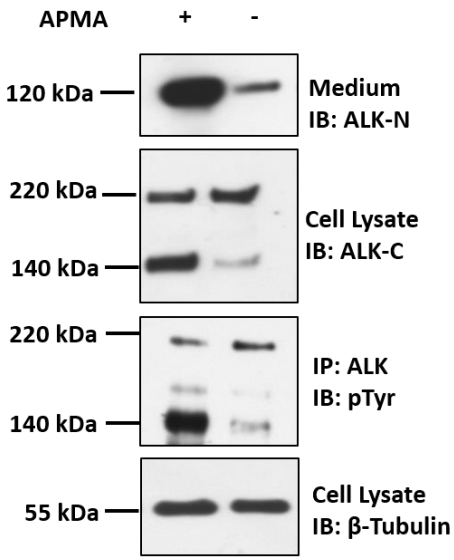


Figure.4.4 140-kDa ALK fragment is highly phosphorylated

A comparison of phosphorylation level in different ALK forms. APMA treatment was applied to induce ALK cleavage. Compared to the full length ALK, the 140-kDa ALK exhibited a much higher level of phosphorylation in western blot.

4.5 Discussion

Here we presented preliminary data for the study of ALK ectodomain shedding. We confirmed that ALK ectodomain shedding was mediated by MMPs and identified the MMP cleavage region in ALK extracellular domain. More importantly, we found that the 140-kDa ALK exhibited much higher phosphorylation levels compared to the full length ALK in non-stimulated conditions. These findings raise the possibility that this ALK fragment may have higher kinase activity compared to its full length form. Similarly, in some neuroblastoma cells, 140-kDa ALK was found overexpressed and had high tyrosine phosphorylation level (George et al., 2008). Recently, different ALK N-terminal truncated forms were found oncogenic either in neuroblastoma or melanoma (Cazes et al., 2013; Okubo et al., 2012; Wiesner et al., 2015). Although these ALK fragments were generated by either alternative splicing or alternative transcription initiation instead of protease cleavage, we could expect that the 140-kDa ALK may share similar oncogenic characteristic, and ectodomain shedding may contribute to the oncogenicity of ALK in some specific cancer types, such as the cancer type with high ALK and MMPs expressions. To further evaluate the hypothesis, we plan to first identify ALK cleavage site and test the oncogenic potential of 140-kDa ALK thoroughly *in vitro*. In addition, we could absolutely quantify the kinase activities of 140-kDa and full length ALKs. If as expected, 140-kDa ALK has stronger kinase activity, we will further investigate the molecular mechanism of 140-kDa ALK over-activation after ectodomain shedding.

4.6 References

Blobel, C. P. (2005). ADAMs: key components in EGFR signalling and development. Nature reviews Molecular cell biology *6*, 32-43.

Cazes, A., Louis-Brennetot, C., Mazot, P., Dingli, F., Lombard, B., Boeva, V., Daveau, R., Cappo, J., Combaret, V., Schleiermacher, G., et al. (2013). Characterization of rearrangements involving the ALK gene reveals a novel truncated form associated with tumor aggressiveness in neuroblastoma. Cancer Res 73, 195-204.

Chen, M. K., and Hung, M. C. (2015). Proteolytic cleavage, trafficking, and functions of nuclear receptor tyrosine kinases. FEBS J *282*, 3693-3721.

Duffy, M. J., McKiernan, E., O'Donovan, N., and McGowan, P. M. (2009). Role of ADAMs in cancer formation and progression. Clin Cancer Res *15*, 1140-1144.

George, R. E., Sanda, T., Hanna, M., Frohling, S., Luther, W., 2nd, Zhang, J., Ahn, Y., Zhou, W., London, W. B., McGrady, P., et al. (2008). Activating mutations in ALK provide a therapeutic target in neuroblastoma. Nature 455, 975-978.

Giri, D. K., Ali-Seyed, M., Li, L. Y., Lee, D. F., Ling, P., Bartholomeusz, G., Wang, S. C., and Hung, M. C. (2005). Endosomal transport of ErbB-2: mechanism for nuclear entry of the cell surface receptor. Mol Cell Biol *25*, 11005-11018.

Hansson, M. D., Rzeznicka, K., Rosenback, M., Hansson, M., and Sirijovski, N. (2008). PCR-mediated deletion of plasmid DNA. Anal Biochem *375*, 373-375.

Hayashida, K., Bartlett, A. H., Chen, Y., and Park, P. W. (2010). Molecular and cellular mechanisms of ectodomain shedding. Anat Rec (Hoboken) *293*, 925-937.

Lafky, J. M., Wilken, J. A., Baron, A. T., and Maihle, N. J. (2008). Clinical implications of the ErbB/epidermal growth factor (EGF) receptor family and its ligands in ovarian cancer. Biochim Biophys Acta *1785*, 232-265.

Lu, P., Takai, K., Weaver, V. M., and Werb, Z. (2011). Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol *3*.

Mazot, P., Cazes, A., Boutterin, M. C., Figueiredo, A., Raynal, V., Combaret, V., Hallberg, B., Palmer, R. H., Delattre, O., Janoueix-Lerosey, I., and Vigny, M. (2011). The constitutive activity of the ALK mutated at positions F1174 or R1275 impairs receptor trafficking. Oncogene *30*, 2017-2025.

Miller, M. A., Meyer, A. S., Beste, M. T., Lasisi, Z., Reddy, S., Jeng, K. W., Chen, C. H., Han, J., Isaacson, K., Griffith, L. G., and Lauffenburger, D. A. (2013). ADAM-10 and -17 regulate endometriotic cell migration via concerted ligand and receptor shedding feedback on kinase signaling. Proceedings of the National Academy of Sciences of the United States of America *110*, E2074-E2083.

Moog-Lutz, C., Degoutin, J., Gouzi, J. Y., Frobert, Y., Brunet-de Carvalho, N., Bureau, J., Creminon, C., and Vigny, M. (2005). Activation and inhibition of anaplastic lymphoma kinase receptor tyrosine kinase by monoclonal antibodies and absence of agonist activity of pleiotrophin. The Journal of biological chemistry *280*, 26039-26048.

Murphy, G. (2008). The ADAMs: signalling scissors in the tumour microenvironment. Nature reviews Cancer *8*, 929-941.

Okubo, J., Takita, J., Chen, Y., Oki, K., Nishimura, R., Kato, M., Sanada, M., Hiwatari, M., Hayashi, Y., Igarashi, T., and Ogawa, S. (2012). Aberrant activation of ALK kinase by a novel truncated form ALK protein in neuroblastoma. Oncogene 31, 4667-4676.

Revillion, F., Lhotellier, V., Hornez, L., Bonneterre, J., and Peyrat, J. P. (2008). ErbB/HER ligands in human breast cancer, and relationships with their receptors, the bio-pathological features and prognosis. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO *19*, 73-80.

Seals, D. F., and Courtneidge, S. A. (2003). The ADAMs family of metalloproteases: multidomain proteins with multiple functions. Genes Dev *17*, 7-30.

Sternlicht, M. D., and Werb, Z. (2001). How matrix metalloproteinases regulate cell behavior. Annual Review of Cell and Developmental Biology *17*, 463.

Visco, V., Bei, R., Moriconi, E., Gianni, W., Kraus, M. H., and Muraro, R. (2000). ErbB2 immune response in breast cancer patients with soluble receptor ectodomain. American Journal of Pathology *156*, 1417.

Weber, S., and Saftig, P. (2012). Ectodomain shedding and ADAMs in development. Development (Cambridge) *139*, 3693-3709.

Wiesner, T., Lee, W., Obenauf, A. C., Ran, L., Murali, R., Zhang, Q. F., Wong, E. W., Hu, W., Scott, S. N., Shah, R. H., et al. (2015). Alternative transcription initiation leads to expression of a novel ALK isoform in cancer. Nature 526, 453-457.

Chapter 5

5 Summary and perspectives

5.1 Summary of study

NUMB is the first reported cell fate determinant, originally characterized in *Drosophila* (Uemura et al., 1989), in which asymmetric distribution of NUMB protein determines the fate of daughter cells during sensory organ precursor division (Rhyu et al., 1994). At the molecular level, NUMB directly binds to and counteracts NOTCH receptor during cell fate determination in this cell lineage (Guo et al., 1996). This is also the pioneer research of asymmetric cell division. Since the NUMB gene is evolutionarily conserved, research on NUMB quickly expanded to encompass advanced animal models. Consistently, NUMB is found to be involved in the development of brain, pancreas and retina in mice (Dho et al., 1999), suggesting that NUMB may generally regulate asymmetric division and cell fate determination during animal development.

The structure of NUMB is conserved from fly to human. As an adapter protein, NUMB has an PTB domain and DPF/NPF motifs, which are responsible for protein-protein interactions and contributing to most known NUMB functions (Bork and Margolis, 1995; Li et al., 1997; Li et al., 1998; Zwahlen et al., 2000). However, the PTB domain and PRR play different roles in NUMB function and have distinct protein binding partners. The PTB domain binds to a number of membrane localized or associated proteins that are crucial for transmembrane signal transduction (Dho et al., 1999; Dho et al., 2006; Gulino et al., 2010), while most known DPF/NPF binding partners are endocytic effector proteins (Salcini et al., 1997; Santolini et al., 2000). This binding pattern implicates a potential role for NUMB in endocytosis regulation. Indeed, NUMB has been shown to regulate the endocytosis and activity of the NOTCH receptor (Barth and Köhler, 2014; McGill et al., 2009; McGill and McGlade, 2003; Yamamoto et al., 2010).

Unlike that in *Drosophila*, the mammalian NUMB gene undergoes alternative splicing to express at least 9 isoforms, of which the four longest have been well studied. The four isoforms are differentiated by the exclusion or inclusion of exon 3 in the PTB domain and exon 9 in the PRR. Even though all four isoforms are involved in development, their functions differ in several animal models, particularly between the PRR^L and PRR^S isoforms as reported in a number of studies (Bani-Yaghoub et al., 2007; Dho et al., 1999a; Dooley et al., 2003; Verdi et al., 1999; Yoshida et al., 2003). In general, PRR^L isoforms promote cell proliferation and PRR^S isoforms stimulate cell differentiation, while their expression patterns vary in different cells at the tissue/organ level. It should be noted that ectopic expression of human PRR^L or PRR^S isoforms results in similar proliferation/differentiation phenotypes in *Drosophila* as well (Toriya et al., 2006). However, the mechanism of this isoform-dependent functional diversity has not been well elucidated. It is only speculated that PRR^L and PRR^S isoforms may recruit distinct effectors that are required for proliferation or differentiation respectively.

Despite the established importance of NUMB in development, current research mainly focuses on studying its function in tumorigenesis. Conventionally, NUMB is defined as a tumor suppressor, due to its negative regulation of the NOTCH receptor, the first reported NUMB binding partner and one of the most frequently deregulated receptors in solid tumors (Espinoza and Miele, 2013; Leong and Gao, 2008). The pivotal role of NOTCH is to control cell fate determination and proliferation at the cellular level. At the tissue level, NOTCH is also involved in angiogenesis and epithelial-mesenchymal transition (Leong and Karsan, 2006). NUMB promotes NOTCH endocytosis and subsequent lysosomal degradation, thus suppressing tumorigenesis. In addition, NUMB also attenuates cell immortality by preventing MDM2 mediated ubiquitination of p53 (Colaluca et al., 2008). On the other hand, NUMB exhibits some oncogenic properties in an isoform dependent manner. PRR^L isoforms, but not PRR^S isoforms, have been shown to correlate with promoted cell proliferation, and the RBM5/6 and RBM10 splicing factors, which are responsible for NUMB exon 9 inclusion or exclusion, regulate the cancer cell proliferation conversely (Bechara et al., 2013; Westhoff et al., 2009).

To further elucidate the complex nature of NUMB function, particularly in tumorigenesis, we systematically investigated the NUMB interactome and defined the binding specificity of NUMB-PTB domain by employing an integrative strategy combining both protein and peptide arrays. The interactome of NUMB-PTB domain was profiled in a protein array that contained 8274 human proteins. Simultaneously, an oriented peptide array library was synthesized and probed with NUMB-PTB domain. The peptide array data were processed using the SMALI program to generate a prediction of NUMB-PTB binding sites in candidate interactors identified in the protein array. Next, the candidate interactors and PTB binding sites were validated via conventional immunoprecipitation, pulldown or protein-peptide binding assays.

It is interesting that a number of RTKs were isolated from the protein array and their binding to NUMB-PTB was confirmed *in vitro* and *in vivo*. Previously, NUMB has been reported to interact with at least three RTKs: EphB2, TrkB and EGFR. These interactions are functional, even though mechanistic demonstration is still lacking (Jiang et al., 2012; Nishimura et al., 2006; Zhou et al., 2011). Some evidence suggests that NUMB likely promotes EGFR endocytosis and lysosomal degradation (Jiang et al., 2012), similar to its regulation of the NOTCH receptor. Our studies exhibit that RTKs are highly enriched in the NUMB interactome and three common oncogenic RTKs, ALK, ErbB2 and Fgr directly interact with NUMB. Because NUMB generally interacts with membrane localized/associated proteins via the PTB domain and recruits endocytic effectors via DPF/NPF motifs, it can be speculated that NUMB may play the same role in connecting RTKs and endocytic effectors, thus promoting RTK endocytosis.

Considering the high accuracy of our SMALI prediction of NUMB-PTB binding sites, we scanned the entire human protein database for isolating more potential NUMB-PTB binding partners. It is intriguing that RTKs and PTPs are highly enriched in the top hit list. These data further increase the possibility that NUMB may act as a universal binding partner and regulator of RTKs.

This hypothesized model is supported by our studies on the interplay between NUMB and ALK. As presented in Chapter 2, we isolated a number of RTKs as NUMB-PTB binding

partners and validated the interactions *in vivo* for ALK, ErbB2 and Fgr. ALK was chosen for subsequent functional studies for two main reasons: (1) ALK is a hot therapeutic target and frequently deregulated in many cancer types. Unlike other oncogenic RTKs, ALK drives the malignant characteristics of cancer cells in many different ways, including gene amplification, overexpression, constitutively active mutations, alternative transcription initiation, and most notably, translocation fusion mutations (Hallberg and Palmer, 2013); At least 22 ALK fusion proteins have been characterized clinically. The mechanism is unclear and it is speculated that the ALK gene is located in an active translocation region in the genome; (2) ALK potentially has less cross-talk with other RTKs. ALK belongs to a small RTK subfamily, consisting of only two members: ALK and LTK, as defined by their unique extracellular domains containing a glycine-rich region (Palmer et al., 2009). ALK is an ideal therapeutic target because it has no physiological significance in adults and only expressed in selected cancer types (Morris et al., 1994). This makes it possible to study ALK in a relatively clean background with less negative impact from functional redundancy or cross-talk among RTKs. We take advantage of a HEK293-based stable cell line with ALK expression for most our studies, which is easily subjected to experimental manipulation. More importantly, we took advantage an ALK stimulating antibody (mimics of ligand) that is highly specific for ALK activation, which greatly facilitates the evaluation of ALK-driven oncogenic behaviors in HEK293/ALK cells.

As predicted, NUMB binds two ALK motifs: G^{1473}GHV<u>NMAF</u>SQ^{1482} and F^{1577}PCGNV<u>NYGY</u>Q^{1587} *in vitro* and *in vivo*. In terms of ALK binding, little difference is found between PTB^{L} and PTB^{S} domains. Both PRR^{L} and PRR^{S} isoforms regulate ALK endocytosis, and direct internalized ALK to distinct destinations during endosomal sorting, as verified by surface biotinylation assays and quantitative fluorescent microscopy. The p72 isoform (PRR^{L}) mainly recycles ALK back to the membrane while the p66 isoform (PRR^{S}) encourages ALK to compartmentalize in the late endosome. Fluorescent microscopy clearly exhibits that ALK co-localizes with different post-endocytic destination markers with p66-NUMB or p72-NUMB expression. The preference for endosomal sorting has different effects on cell proliferation and anchorage-dependent cell growth that are driven by ALK activation. In particular, the p72 isoform assists in sustaining the active status of the ALK and MAPK proliferation signal pathway.

Conversely, the p66 isoform promotes ALK degradation, thus attenuating ALK activity and MAPK signaling.

Our findings provide the first evidence for the differential functions of NUMB isoforms in endosomal sorting and the regulation of RTK activity. NUMB generally promotes the endocytosis of ALK and further determines the destination of internalized ALK in an isoform-dependent manner. The p72 isoform recycles ALK back to the membrane to sustain ALK activity, thus being potentially oncogenic. In contrast, the p66 isoform mainly directs ALK to the late endosome for lysosomal degradation. Considering the functional pattern of NUMB proteins, it is possible that the NUMB isoforms may act in a similar manner for all NUMB-binding receptors, including RTKs. This is consistent with previous reports that the NUMB PRR^{L} isoform correlates with malignant behaviors in a number of cancer types (Bechara et al., 2013; Misquitta-Ali et al., 2011; Westhoff et al., 2009).

5.2 NUMB isoforms, endosomal sorting and cancer diagnosis

NUMB is defined as an endocytic adaptor. As introduced above, the PTB domain binds to receptors on the membrane while the DPF/NPF recruits endocytic effector proteins. This functional pattern raises a possibility that NUMB may act as a universal bi-directional regulator of receptors by regulating endocytosis and subsequent endosomal sorting of receptors. It has been reported that NUMB is involved in both recycling and endosomal compartmentalization of NOTCH (McGill et al., 2009), and NUMB alternative splicing could increase the activity of NOTCH and its downstream signaling (Bechara et al., 2013; Westhoff et al., 2009). This is in agreement with our model, and it is likely that NUMB regulates NOTCH in the same way.

An immediate question to address is how PRR^{L} and PPR^{S} determine the direction of endosomal sorting. It can be speculated that PRR^{L} and PPR^{S} may recruit different sets of endocytic effectors that in turn decide the destination of internalized proteins. But this is challenged by a mass spectrometry based study that compares the interactome between p72 and p66 isoforms. It is reported that both isoforms bind to the same few endocytic proteins, but with different binding affinities (Krieger et al., 2013). However, only a limited number

of binding partners were identified in this study and we cannot exclude the possibility that PRR^L and PPR^S recruit different endocytic proteins. In our study, even with the overexpression of the two isoforms, the differences in cell proliferation and anchorage-dependent growth are only moderate. This suggests that p72 and p66 isoforms may not play absolutely opposing roles in ALK endosomal sorting; instead, they likely just prefer the recycling or late endosomal compartmentalization of ALK, respectively. To further characterize the function of NUMB isoforms in endosomal sorting, a structural analysis and systematic interactome profile are necessary. In addition, the strategy employed in our studies may be repeated for the study of NUMB and other RTKs, but it will be important to establish a proper cell line/system to carry out these experiments.

Based on our studies and previous reports, the p72 isoform may contribute to the oncogenicity of NUMB by sustaining the activity of membrane localized receptors. On the other hand, NUMB is able to suppress tumor growth in parallel pathways, either by promoting receptor endocytosis and subsequent degradation, or by stabilizing p53 to attenuate cell immortality. Therefore, simultaneously identifying of the status of p53, receptors, and NUMB alternative splicing forms may better predict the prognosis of specific tumor types and guide the design of adjuvant therapy. We have scheduled a computational analysis for evaluating p53 levels, NOTCH/ALK/ErbB2 expression levels and NUMB alternative splicing forms from published cancer cell microarray data. For select cell lines with distinct fingerprints, we will thoroughly characterize their oncogenic behaviors and responses to target therapeutics.

5.3 References

Bani-Yaghoub, M., Kubu, C. J., Cowling, R., Rochira, J., Nikopoulos, G. N., Bellum, S., and Verdi, J. M. (2007). A switch in numb isoforms is a critical step in cortical development. Developmental Dynamics *236*, 696-705.

Barth, J. M. I., and Köhler, K. (2014). How to take autophagy and endocytosis up a notch. BioMed Research International *2014*.

Bechara, E. G., Sebestyen, E., Bernardis, I., Eyras, E., and Valcarcel, J. (2013). RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. Molecular cell *52*, 720-733.

Bork, P., and Margolis, B. (1995). A phosphotyrosine interaction domain. Cell *80*, 693-694.

Colaluca, I. N., Tosoni, D., Nuciforo, P., Senic-Matuglia, F., Galimberti, V., Viale, G., Pece, S., and Di Fiore, P. P. (2008). NUMB controls p53 tumour suppressor activity. Nature *451*, 76-80.

Dho, S. E., French, M. B., Woods, S. A., and McGlade, C. J. (1999). Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. J Biol Chem *274*, 33097-33104.

Dho, S. E., Trejo, J., Siderovski, D. P., and McGlade, C. J. (2006). Dynamic regulation of mammalian numb by G protein-coupled receptors and protein kinase C activation: Structural determinants of numb association with the cortical membrane. Molecular biology of the cell *17*, 4142-4155.

Dooley, C. M., James, J., McGlade, C. J., and Ahmad, I. (2003). Involvement of numb in vertebrate retinal development: Evidence for multiple roles of numb in neural differentiation and maturation. Journal of Neurobiology *54*, 313-325.

Espinoza, I., and Miele, L. (2013). Notch inhibitors for cancer treatment. Pharmacology and Therapeutics *139*, 95-110.

Gulino, A., Di Marcotullio, L., and Screpanti, I. (2010). The multiple functions of Numb. Exp Cell Res *316*, 900-906.

Guo, M., Jan, L. Y., and Jan, Y. N. (1996). Control of daughter cell fates during asymmetric division: Interaction of Numb and Notch. Neuron *17*, 27-41.

Hallberg, B., and Palmer, R. H. (2013). Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. Nature reviews Cancer *13*, 685-700.

Jiang, X., Xing, H., Kim, T. M., Jung, Y., Huang, W., Yang, H. W., Song, S., Park, P. J., Carroll, R. S., and Johnson, M. D. (2012). Numb regulates glioma stem cell fate and growth by altering epidermal growth factor receptor and Skp1-Cullin-F-box ubiquitin ligase activity. Stem cells *30*, 1313-1326.

Leong, K. G., and Gao, W. Q. (2008). The Notch pathway in prostate development and cancer. Differentiation *76*, 699-716.

Leong, K. G., and Karsan, A. (2006). Recent insights into the role of Notch signaling in tumorigenesis. Blood *107*, 2223-2233.

Li, S. C., Sonoyang, Z., Vincent, S. J. F., Zwahlen, C., Wiley, S., Cantley, L., Kay, L. E., Forman-Kay, J., and Pawson, T. (1997). High-affinity binding of the Drosophila Numb phosphotyrosine-binding domain to peptides containing a Gly-Pro-(p)Tyr motif. Proceedings of the National Academy of Sciences of the United States of America *94*, 7204-7209.

Li, S. C., Zwahlen, C., Vincent, S. J., McGlade, C. J., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (1998). Structure of a Numb PTB domain-peptide complex suggests a basis for diverse binding specificity. Nature structural biology *5*, 1075-1083.

McGill, M. A., Dho, S. E., Weinmaster, G., and McGlade, C. J. (2009). Numb regulates post-endocytic trafficking and degradation of notch1. Journal of Biological Chemistry *284*, 26427-26438.

McGill, M. A., and McGlade, C. J. (2003). Mammalian numb proteins promote Notch1 receptor ubiquitination and degradation of the Notch1 intracellular domain. J Biol Chem *278*, 23196-23203.

Morris, S. W., Kirstein, M. N., Valentine, M. B., Dittmer, K. G., Shapiro, D. N., Saltman, D. L., and Look, A. T. (1994). Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science *263*, 1281-1284.

Nishimura, T., Yamaguchi, T., Tokunaga, A., Hara, A., Hamaguchi, T., Kato, K., Iwamatsu, A., Okano, H., and Kaibuchi, K. (2006). Role of numb in dendritic spine development with a Cdc42 GEF intersectin and EphB2. Molecular biology of the cell *17*, 1273-1285.

Palmer, R. H., Vernersson, E., Grabbe, C., and Hallberg, B. (2009). Anaplastic lymphoma kinase: Signalling in development and disease. Biochemical Journal *420*, 345-361.

Rhyu, M. S., Jan, L. Y., and Jan, Y. N. (1994). Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells. Cell *76*, 477-491.

Salcini, A. E., Confalonieri, S., Doria, M., Santolini, E., Tassi, E., Minenkova, O., Cesareni, G., Pelicci, P. G., and Di Fiore, P. P. (1997). Binding specificity and in vivo targets of the EH domain, a novel protein-protein interaction module. Genes Dev *11*, 2239-2249.

Santolini, E., Puri, C., Salcini, A. E., Gagliani, M. C., Pelicci, P. G., Tacchetti, C., and Di Fiore, P. P. (2000). Numb is an endocytic protein. The Journal of cell biology *151*, 1345-1352.

Toriya, M., Tokunaga, A., Sawamoto, K., Nakao, K., and Okano, H. (2006). Distinct functions of human numb isoforms revealed by misexpression in the neural stem cell lineage in the Drosophila larval brain. Developmental Neuroscience *28*, 142-155.

Uemura, T., Shepherd, S., Ackerman, L., Jan, L. Y., and Jan, Y. N. (1989). numb, a gene required in determination of cell fate during sensory organ formation in Drosophila embryos. Cell *58*, 349-360.

Verdi, J. M., Bashirullah, A., Goldhawk, D. E., Kubu, C. J., Jamali, M., Meakin, S. O., and Lipshitz, H. D. (1999). Distinct human NUMB isoforms regulate differentiation vs. proliferation in the neuronal lineage. Proceedings of the National Academy of Sciences of the United States of America *96*, 10472-10476.

Westhoff, B., Colaluca, I. N., D'Ario, G., Donzelli, M., Tosoni, D., Volorio, S., Pelosi, G., Spaggiari, L., Mazzarol, G., Viale, G., *et al.* (2009). Alterations of the Notch pathway in lung cancer. Proc Natl Acad Sci U S A *106*, 22293-22298.

Yamamoto, S., Charng, W. L., and Bellen, H. J. (2010). Endocytosis and intracellular trafficking of notch and its ligands. In Current Topics in Developmental Biology, pp. 165-200.

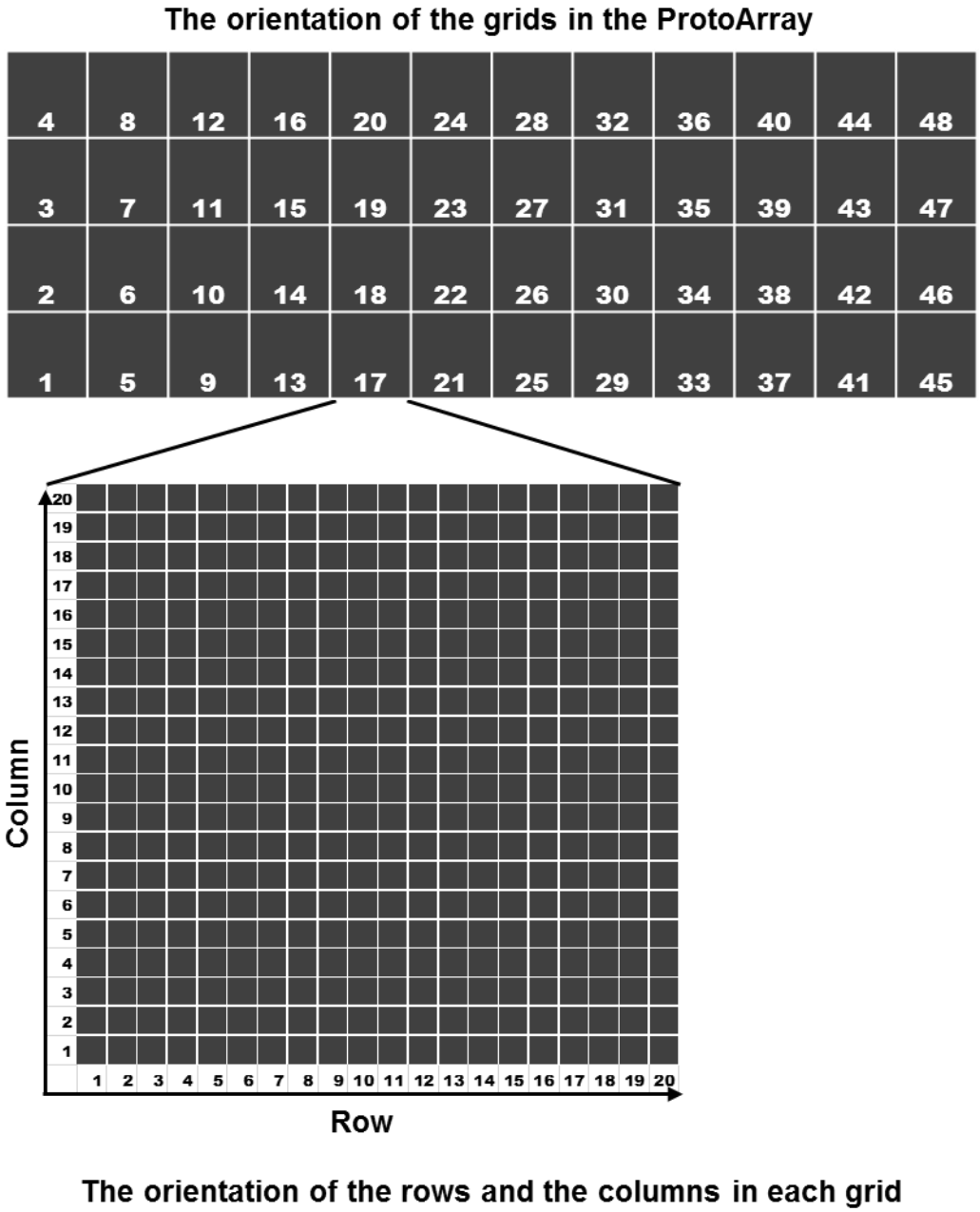
Yoshida, T., Tokunaga, A., Nakao, K., and Okano, H. (2003). Distinct expression patterns of splicing isoforms of mNumb in the endocrine lineage of developing pancreas. Differentiation *71*, 486-495.

Zhou, P., Alfaro, J., Chang, E. H., Zhao, X., Porcionatto, M., and Segal, R. A. (2011). Numb links extracellular cues to intracellular polarity machinery to promote chemotaxis. Dev Cell *20*, 610-622.

Zwahlen, C., Li, S. C., Kay, L. E., Pawson, T., and Forman-Kay, J. D. (2000). Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. The EMBO journal *19*, 1505-1515.

Supplemental Data

S1 ProtoArray grid design



S2 ProtoArray quantification of NUMB-PTB domain (Z-Score>3.0)

Rank	Uniprot Name	Grid	Row	Column	Signal	Z-Score
1	PVRL3_HUMAN	3	18	5	36315.3	30.74933
2	PVRL3_HUMAN	3	18	6	33886.8	28.66454
3	GLYG2_HUMAN	1	17	20	30656.6	25.89145
4	SLAI2_HUMAN	48	15	10	27682.6	23.33829
5	SLAI2_HUMAN	48	15	9	23388.9	19.6522
6	CJ091_HUMAN	17	18	3	22245.8	18.67084
7	CJ091_HUMAN	17	18	4	21979	18.44183
8	PIM1_HUMAN	1	19	20	20991.3	17.59388
9	LCK_HUMAN	39	19	19	19938.7	16.69021
10	LCK_HUMAN	39	19	20	19769.6	16.54503
11	M3K2_HUMAN	42	19	18	17040.1	14.2018
12	M3K2_HUMAN	42	19	17	16796.4	13.9926
13	FGFR2_HUMAN	48	19	18	16583.1	13.80947
14	FGFR4_HUMAN	47	19	20	15525.1	12.90122
15	TOPK_HUMAN	27	19	16	15383	12.77921
16	CDK5_HUMAN	44	19	12	15223.1	12.64197
17	FGFR2_HUMAN	48	19	17	15196.7	12.61931
18	TOPK_HUMAN	27	19	15	14900	12.3646
19	FGFR4_HUMAN	47	19	19	14785.2	12.26599
20	FYN_HUMAN	3	19	19	14654	12.15338
21	PIM1_HUMAN	1	19	19	14214.7	11.77621
22	FYN_HUMAN	3	19	20	14207.8	11.7703
23	ALK_HUMAN	41	19	18	13745.1	11.37306
24	GLYG2_HUMAN	1	17	19	13543.5	11.19998
25	Q96E28_HUMAN	48	12	20	13396.9	11.07417
26	ALK_HUMAN	41	19	17	13103.2	10.82204
27	FLT3_HUMAN	3	19	13	12455	10.26558
28	KPCB_HUMAN	3	19	9	12054.5	9.92171
29	Q96E28_HUMAN	48	12	19	11980.4	9.85811
30	FLT3_HUMAN	3	19	14	11912	9.79937
31	FGR_HUMAN	30	19	20	11650.1	9.57456
32	FGR_HUMAN	30	19	19	10870.2	8.90499
33	CDK5_HUMAN	44	19	11	10537.2	8.61912
34	KPCB_HUMAN	3	19	10	10188.4	8.31969
35	EMC9_HUMAN	23	16	3	9668.02	7.87296
36	CDK2_HUMAN	42	19	13	9578.64	7.79623
37	FGFR3_HUMAN	9	19	11	9517.92	7.7441

38	CSK21_HUMAN	5	19	13	9484.94	7.71578
39	CUED1_HUMAN	47	5	19	9330.07	7.58283
40	LIMC1_HUMAN	1	7	20	9319.14	7.57345
41	IGF1R_HUMAN	41	19	15	8877.21	7.19406
42	HCK_HUMAN	35	19	10	8868.84	7.18687
43	LYN_HUMAN	11	19	9	8780.18	7.11076
44	RIPK2_HUMAN	3	19	18	8775.8	7.107
45	FLT3_HUMAN	35	19	13	8752.18	7.08672
46	KPCD_HUMAN	19	15	5	8748.06	7.08318
47	AKT1_HUMAN	44	19	10	8720.24	7.0593
48	CUED1_HUMAN	47	5	20	8713.85	7.05381
49	FGFR3_HUMAN	9	19	12	8641.57	6.99177
50	CDK2_HUMAN	42	19	14	8631.38	6.98301
51	BTK_HUMAN	4	19	9	8459.87	6.83578
52	PADI4_HUMAN	1	1	19	8441.76	6.82023
53	RIPK2_HUMAN	3	19	17	8384.91	6.77142
54	YES_HUMAN	25	19	14	8299.2	6.69784
55	AKT1_HUMAN	44	19	9	8285.45	6.68603
56	CSK21_HUMAN	5	19	14	8188.7	6.60298
57	ROS1_HUMAN	20	19	10	8188.17	6.60253
58	TBK1_HUMAN	37	19	13	8170.14	6.58705
59	FLT3_HUMAN	35	19	14	8168.28	6.58544
60	BTK_HUMAN	4	19	10	8154.94	6.57399
61	NTKL_HUMAN	48	13	4	8128.75	6.55151
62	SRPK1_HUMAN	1	19	14	8119.07	6.54321
63	ROS1_HUMAN	20	19	9	8091.6	6.51962
64	CSF1R_HUMAN	37	19	17	8076.69	6.50682
65	PPID_HUMAN	42	18	9	8062.23	6.49441
66	YES_HUMAN	25	19	13	8038.5	6.47403
67	PPID_HUMAN	42	18	10	7961	6.4075
68	KPCG_HUMAN	1	19	11	7897.96	6.35338
69	MP2K1_HUMAN	2	19	11	7893.63	6.34967
70	NTKL_HUMAN	48	13	3	7868.84	6.32838
71	IGF1R_HUMAN	41	19	16	7806.25	6.27465
72	KPCG_HUMAN	1	19	12	7755.49	6.23107
73	HCK_HUMAN	35	19	9	7714.97	6.19629
74	JAK2_HUMAN	4	19	11	7690.65	6.17541
75	SRPK1_HUMAN	1	19	13	7610.42	6.10653
76	PDPK1_HUMAN	10	19	12	7579.43	6.07993

77	PDPK1_HUMAN	10	19	11	7506.73	6.01752
78	IKKA_HUMAN	47	19	13	7466.15	5.98267
79	TBK1_HUMAN	37	19	14	7426.09	5.94828
80	EMC8_HUMAN	1	3	13	7316.67	5.85435
81	FBX38_HUMAN	35	16	15	7316.45	5.85416
82	IKKA_HUMAN	47	19	14	7291.03	5.83234
83	EMC8_HUMAN	1	3	14	7246.75	5.79432
84	LYN_HUMAN	11	19	10	7206.27	5.75957
85	QKI_HUMAN	45	5	4	7175.25	5.73294
86	LYN_HUMAN	45	19	11	7147.52	5.70914
87	BLK_HUMAN	31	19	20	7107.49	5.67477
88	TOPK_HUMAN	42	9	19	7046.54	5.62245
89	MP2K1_HUMAN	2	19	12	7019.51	5.59924
90	TOPK_HUMAN	42	9	20	6999.5	5.58206
91	JAK2_HUMAN	4	19	12	6980.62	5.56585
92	PADI4_HUMAN	1	1	20	6959.65	5.54785
93	BLK_HUMAN	31	19	19	6818.67	5.42682
94	ST17A_HUMAN	46	19	17	6698.89	5.32399
95	LYN_HUMAN	45	19	12	6626.09	5.26149
96	HCK_HUMAN	44	16	6	6508.78	5.16078
97	EPHB4_HUMAN	27	19	19	6491.82	5.14623
98	CT55_HUMAN	45	7	3	6358.62	5.03188
99	SGK1_HUMAN	21	11	2	6313.66	4.99328
100	ST17A_HUMAN	46	19	18	6187.43	4.88491
101	Q8IY52_HUMAN	18	17	8	6117.21	4.82463
102	RBP10_HUMAN	45	19	20	6097.39	4.80761
103	CT55_HUMAN	45	7	4	6079.7	4.79243
104	KS6A1_HUMAN	24	19	9	6048.34	4.7655
105	KPCT_HUMAN	19	19	14	6027.71	4.74779
106	KS6A1_HUMAN	24	19	10	6003.57	4.72707
107	CSF1R_HUMAN	37	19	18	5900.46	4.63855
108	SDCG3_HUMAN	26	15	5	5883.08	4.62362
109	Q8IY52_HUMAN	18	17	7	5785.89	4.54019
110	HCK_HUMAN	44	16	5	5782.46	4.53724
111	PTK6_HUMAN	32	19	10	5728.43	4.49086
112	SGK1_HUMAN	21	11	1	5721.5	4.48492
113	RPC6_HUMAN	44	14	13	5660.13	4.43223
114	TGFR2_HUMAN	35	14	13	5651.01	4.42439
115	KPCA_HUMAN	1	19	18	5620.63	4.39831

116	AR6P4_HUMAN	25	6	18	5593.06	4.37465
117	KPCD_HUMAN	5	19	11	5580.17	4.36358
118	KPCT_HUMAN	19	19	13	5508.64	4.30217
119	KPCD_HUMAN	5	19	12	5490.67	4.28675
120	NUSAP_HUMAN	18	5	18	5488.44	4.28483
121	LMX1B_HUMAN	42	11	7	5484.62	4.28156
122	PAK4_HUMAN	32	19	14	5453.11	4.25451
123	ILKAP_HUMAN	36	11	7	5434.85	4.23882
124	SDCG3_HUMAN	26	15	6	5385.02	4.19605
125	KPCA_HUMAN	1	19	17	5373.54	4.18619
126	LMX1B_HUMAN	42	11	8	5315.51	4.13637
127	TM1L2_HUMAN	8	16	1	5285.32	4.11045
128	PAK4_HUMAN	32	19	13	5240.19	4.07171
129	UFO_HUMAN	47	19	18	5229.03	4.06213
130	TM169_HUMAN	23	16	2	5192.73	4.03097
131	TGFR2_HUMAN	35	14	14	5167.85	4.00961
132	ILKAP_HUMAN	36	11	8	5128.1	3.97548
133	SRC_HUMAN	7	19	19	5120.55	3.969
134	GBA2_HUMAN	44	11	8	5113.66	3.96309
135	PDE4A_HUMAN	36	4	7	5050.26	3.90866
136	MYL6B_HUMAN	2	15	8	5047.82	3.90657
137	PCCA_HUMAN	39	4	15	5028.7	3.89015
138	PTK6_HUMAN	32	19	9	5007.69	3.87212
139	LIMC1_HUMAN	1	7	19	4981.93	3.85
140	MP2K1_HUMAN	35	19	15	4978.09	3.84671
141	UBXN6_HUMAN	37	5	3	4951.69	3.82404
142	PAK3_HUMAN	19	19	18	4939.15	3.81327
143	EPHA8_HUMAN	48	19	11	4936.82	3.81128
144	BMX_HUMAN	17	19	15	4932.6	3.80765
145	QKI_HUMAN	45	5	3	4919.16	3.79611
146	RM19_HUMAN	46	17	17	4902.87	3.78213
147	MYL6B_HUMAN	2	15	7	4898.98	3.77878
148	M4K2_HUMAN	33	19	19	4895.71	3.77598
149	CK074_HUMAN	7	19	5	4878.44	3.76115
150	ERBB2_HUMAN	17	19	11	4795.76	3.69018
151	ARFG1_HUMAN	36	6	3	4786.86	3.68253
152	ANS4B_HUMAN	19	14	2	4786.81	3.68249
153	SRC_HUMAN	7	19	20	4783.57	3.67971
154	PAK3_HUMAN	19	19	17	4762.73	3.66182

155	Q8TB46_HUMAN	18	2	12	4761.44	3.66071
156	UFO_HUMAN	47	19	17	4743.09	3.64496
157	SMTL2_HUMAN	45	15	3	4727.37	3.63146
158	RM19_HUMAN	46	17	18	4703.64	3.61109
159	TM1L2_HUMAN	8	16	2	4688.19	3.59783
160	L7RTI5_HUMAN	4	19	13	4639.4	3.55594
161	M4K2_HUMAN	33	19	20	4633.61	3.55097
162	UBXN6_HUMAN	37	5	4	4632.58	3.55009
163	BMX_HUMAN	17	19	16	4591.67	3.51497
164	TEC_HUMAN	25	19	20	4560.56	3.48826
165	YES_HUMAN	48	15	17	4558.41	3.48641
166	BEGIN_HUMAN	37	8	5	4548.89	3.47824
167	ARFG1_HUMAN	36	6	4	4544.26	3.47427
168	EPHB4_HUMAN	27	19	20	4531.18	3.46304
169	Q8TB46_HUMAN	18	2	11	4522.1	3.45524
170	ERBB2_HUMAN	17	19	12	4515.15	3.44927
171	M3K3_HUMAN	35	19	11	4482.24	3.42102
172	NTRK1_HUMAN	39	19	11	4458.61	3.40074
173	IRS1_HUMAN	19	7	19	4451.18	3.39435
174	SMTL2_HUMAN	45	15	4	4448.79	3.39231
175	M3K11_HUMAN	33	19	16	4446.09	3.38998
176	PCCA_HUMAN	39	4	16	4405.68	3.35529
177	NUAK1_HUMAN	6	19	18	4389.85	3.3417
178	SNAG_HUMAN	11	11	12	4382.94	3.33577
179	M3K11_HUMAN	33	19	15	4373.26	3.32746
180	RBP10_HUMAN	45	19	19	4364.82	3.32022
181	MET_HUMAN	13	19	20	4363.51	3.31909
182	TEC_HUMAN	25	19	19	4350.54	3.30796
183	DAPK2_HUMAN	33	19	14	4318.56	3.2805
184	IRAK4_HUMAN	16	19	17	4295.26	3.2605
185	Q8WUK8_HUMAN	5	4	9	4284.99	3.25168
186	FRK_HUMAN	46	19	20	4274.22	3.24244
187	PCKGC_HUMAN	46	17	11	4273.91	3.24217
188	TMOD2_HUMAN	12	14	9	4259.88	3.23013
189	CK074_HUMAN	7	19	6	4250.07	3.22171
190	YES_HUMAN	48	15	18	4245.22	3.21754
191	JAK3_HUMAN	43	19	17	4241.14	3.21404
192	FRK_HUMAN	46	19	19	4240.98	3.2139
193	MP2K1_HUMAN	35	19	16	4238.36	3.21165

194	MAPK3_HUMAN	30	19	15	4235.18	3.20892
195	MET_HUMAN	13	19	19	4220.96	3.19672
196	HOIL1_HUMAN	19	18	2	4218.22	3.19437
197	IRS1_HUMAN	19	7	20	4192.95	3.17267
198	NTRK1_HUMAN	39	19	12	4188.01	3.16843
199	AURKB_HUMAN	31	19	11	4162.79	3.14678
200	IRAK4_HUMAN	16	19	18	4149.08	3.13501
201	MYL1_HUMAN	1	10	17	4135.88	3.12368
202	TMOD2_HUMAN	12	14	10	4132.82	3.12104
203	CACO1_HUMAN	22	19	2	4128.69	3.1175
204	CPLX2_HUMAN	2	5	11	4121.33	3.11119
205	MYL1_HUMAN	1	10	18	4113.2	3.1042
206	JAK3_HUMAN	43	19	18	4081.93	3.07736
207	B3KVH4_HUMAN	10	11	11	4068.79	3.06608
208	EPHA8_HUMAN	48	19	12	4067.12	3.06464
209	NUAK1_HUMAN	6	19	17	4055.6	3.05475
210	TOM1_HUMAN	18	15	12	4046.53	3.04697
211	AURKB_HUMAN	31	19	12	4039.26	3.04073
212	ABL2_HUMAN	28	19	14	4029.12	3.03203
213	TOM1_HUMAN	18	15	11	4013.05	3.01822
214	BEGIN_HUMAN	37	8	6	4005.81	3.01201
215	RPC6_HUMAN	44	14	14	3999.14	3.00628
216	KCC2B_HUMAN	42	19	16	3994.02	3.00189

S3 NUMB interactome

#	Uniprot Name	Gene names
1	ABL2_HUMAN	ABL2 ABLL ARG
2	AKT1_HUMAN	AKT1 PKB RAC
3	ALK_HUMAN	ALK
4	ANS4B_HUMAN	ANKS4B HARP
5	AR6P4_HUMAN	ARL6IP4
6	ARFG1_HUMAN	ARFGAP1 ARF1GAP
7	AURKB_HUMAN	AURKB AIK2 AIM1 AIRK2 ARK2 STK1 STK12 STK5
8	B3KVH4_HUMAN	
9	BEGIN_HUMAN	BEGAIN KIAA1446
10	BLK_HUMAN	BLK
11	BMX_HUMAN	BMX
12	BTK_HUMAN	BTK AGMX1 ATK BPK
13	CACO1_HUMAN	CALCOCO1 KIAA1536 PP13275 UNQ2436/PRO4996
14	CDK2_HUMAN	CDK2 CDKN2
15	CDK5_HUMAN	CDK5 CDKN5
16	CJ091_HUMAN	C10orf91
17	CK074_HUMAN	C11orf74
18	CPLX2_HUMAN	CPLX2
19	CSF1R_HUMAN	CSF1R FMS
20	CSK21_HUMAN	CSNK2A1 CK2A1
21	CT55_HUMAN	CT55 CXorf48
22	CUED1_HUMAN	CUEDC1
23	DAPK2_HUMAN	DAPK2
24	EMC8_HUMAN	EMC8 C16orf2 C16orf4 COX4AL COX4NB FAM158B NOC4
25	EMC9_HUMAN	EMC9 C14orf122 FAM158A CGI-112
26	EPHA8_HUMAN	EPHA8 EEK HEK3 KIAA1459
27	EPHB4_HUMAN	EPHB4 HTK MYK1 TYRO11
28	ERBB2_HUMAN	ERBB2 HER2 MLN19 NEU NGL
29	FBX38_HUMAN	FBXO38 SP329
30	FGFR2_HUMAN	FGFR2 BEK KGFR KSAM
31	FGFR3_HUMAN	FGFR3 JTK4
32	FGFR4_HUMAN	FGFR4 JTK2 TKF
33	FGR_HUMAN	FGR SRC2
34	FLT3_HUMAN	FLT3 CD135 FLK2 STK1
35	FRK_HUMAN	FRK PTK5 RAK
36	FYN_HUMAN	FYN

37	GBA2_HUMAN	GBA2 KIAA1605 AD035
38	GLYG2_HUMAN	GYG2
39	HCK_HUMAN	HCK
40	HOIL1_HUMAN	RBCK1 C20orf18 RNF54 UBCE7IP3 XAP3 XAP4
41	IGF1R_HUMAN	IGF1R
42	IKKA_HUMAN	CHUK IKKA TCF16
43	ILKAP_HUMAN	ILKAP
44	IRAK4_HUMAN	IRAK4
45	IRS1_HUMAN	IRS1
46	JAK2_HUMAN	JAK2
47	JAK3_HUMAN	JAK3
48	KCC2B_HUMAN	CAMK2B CAM2 CAMK2 CAMKB
49	KPCA_HUMAN	PRKCA PKCA PRKACA
50	KPCB_HUMAN	PRKCB PKCB PRKCB1
51	KPCD_HUMAN	PRKCD
52	KPCG_HUMAN	PRKCG PKCG
53	KPCT_HUMAN	PRKCQ PRKCT
54	KS6A1_HUMAN	RPS6KA1 MAPKAPK1A RSK1
55	L7RTI5_HUMAN	PRKCE
56	LCK_HUMAN	LCK
57	LIMC1_HUMAN	LIMCH1 KIAA1102
58	LMX1B_HUMAN	LMX1B
59	LYN_HUMAN	LYN JTK8
60	M3K11_HUMAN	MAP3K11 MLK3 PTK1 SPRK
61	M3K2_HUMAN	MAP3K2 MAPKKK2 MEKK2
62	M3K3_HUMAN	MAP3K3 MAPKKK3 MEKK3
63	M4K2_HUMAN	MAP4K2 GCK RAB8IP
64	MAPK3_HUMAN	MAPKAPK3
65	MET_HUMAN	MET
66	MP2K1_HUMAN	MAP2K1 MEK1 PRKMK1
67	MYL1_HUMAN	MYL1
68	MYL6B_HUMAN	MYL6B MLC1SA
69	NTKL_HUMAN	SCYL1 CVAK90 GKLP NTKL TAPK TEIF TRAP HT019
70	NTRK1_HUMAN	NTRK1 MTC TRK TRKA
71	NUAK1_HUMAN	NUAK1 ARK5 KIAA0537 OMPHK1
72	NUSAP_HUMAN	NUSAP1 ANKT BM-037 PRO0310
73	PADI4_HUMAN	PADI4 PAD4 PADI5 PDI5
74	PAK3_HUMAN	PAK3 OPHN3
75	PAK4_HUMAN	PAK4 KIAA1142

76	PCCA_HUMAN	PCCA
77	PCKGC_HUMAN	PCK1 PEPCK1
78	PDE4A_HUMAN	PDE4A DPDE2
79	PDPK1_HUMAN	PDPK1 PDK1
80	PIM1_HUMAN	PIM1
81	PPID_HUMAN	PPID CYP40 CYPD
82	PTK6_HUMAN	PTK6 BRK
83	PVRL3_HUMAN	PVRL3 PRR3
84	Q8IY52_HUMAN	DYNC1I1
85	Q8TB46_HUMAN	ANKRD50
86	Q8WUK8_HUMAN	STAC
87	Q96E28_HUMAN	SULF1
88	QKI_HUMAN	QKI HKQ
89	RBP10_HUMAN	RANBP10 KIAA1464
90	RIPK2_HUMAN	RIPK2 CARDIAK RICK RIP2 UNQ277/PRO314/PRO34092
91	RM19_HUMAN	MRPL19 KIAA0104 MRPL15
92	ROS1_HUMAN	ROS1 MCF3 ROS
93	RPC6_HUMAN	POLR3F
94	SDCG3_HUMAN	SDCCAG3
95	SGK1_HUMAN	SGK1 SGK
96	SLAI2_HUMAN	SLAIN2 KIAA1458
97	SMTL2_HUMAN	SMTNL2
98	SNAG_HUMAN	NAPG SNAPG
99	SRC_HUMAN	SRC SRC1
100	SRPK1_HUMAN	SRPK1
101	ST17A_HUMAN	STK17A DRAK1
102	TBK1_HUMAN	TBK1 NAK
103	TEC_HUMAN	TEC PSCTK4
104	TGFR2_HUMAN	TGFBR2
105	TM169_HUMAN	TMEM169
106	TM1L2_HUMAN	TOM1L2
107	TMOD2_HUMAN	TMOD2 NTMOD
108	TOM1_HUMAN	TOM1
109	TOPK_HUMAN	PBK TOPK
110	UBXN6_HUMAN	UBXN6 UBXD1 UBXDC2
111	UFO_HUMAN	AXL UFO
112	YES_HUMAN	YES1 YES

S4 NUMB-PTB peptide array quantification

Position	Intensity	Sequences	SMALI Score	Protein Name	Z-Score
1:1	0.347763	VLEENLYFQGS	0.4818	VGFR3_HUMAN	1.46173
1:2	0.300605	GVQINEEFCQR	0.6127	VGFR3_HUMAN	1.46173
1:3	0.662353	KLASNYVYEVM	0.7411	TEC_HUMAN	3.30796
1:4	0.671989	VHDANTLYIFA	0.346	TEC_HUMAN	3.30796
1:5	0.36597	DKYGNEGYIPS	1.455	TEC_HUMAN	3.30796
1:6	0.387521	IFKEHNFSFYP	0.481	TEC_HUMAN	3.30796
1:7	0.97945	FPRENLEFGKV	0.5042	FLT3_HUMAN	10.26558
1:8	0.953695	GSSDNEYFYVD	0.8388	FLT3_HUMAN	10.26558
1:9	0.525919	TMFPNMDYDII	0.8348	CUED1_HUMAN	7.58283
1:10	0.722152	LFLQNEEFMKE	0.5733	CUED1_HUMAN	7.58283
1:11	0.771759	ELQRNRDFLLA	0.362	CUED1_HUMAN	7.58283
1:12	1.281694	VDPDNVTFCVL	0.7521	GA45G_HUMAN	-0.00226
1:13	0.823835	RDVHNLDYYKK	0.623	FGFR3_HUMAN	7.7441
2:1	1.382014	PMHSNHYYQTV	0.3074	C1QT1_HUMAN	-0.31358
2:2	0.878876	GERENAIFSEE	0.543	C1QT1_HUMAN	-0.31358
2:3	0.86461	ALAINKSFDQR	0.9228	3MG_HUMAN	1.03914
2:4	0.841948	IHTENKLYLVF	0.4167	CDK2_HUMAN	7.79623
2:5	1.309699	PTTENLYFQGA	0.5379	CDK5_HUMAN	8.61912
2:6	2.698102	EKYKNSKYHGV	0.8614	CDK2_HUMAN	7.79623
2:7	1.87173	NPICNDHYRSV	0.5912	TOPK_HUMAN	12.3646
2:8	0.842733	TLATNDIYCQG	0.2427	GLYG2_HUMAN	11.19998
2:9	0.980089	QGLLNSFFRNW	0.5015	GLYG2_HUMAN	11.19998
2:10	1.114231	MKPWNYKYNPQ	0.6042	GLYG2_HUMAN	11.19998
2:11	1.452584	NLSSNTMYTYS	0.3743	GLYG2_HUMAN	11.19998
2:12	1.025716	DNAKNSLYLQM	0.3037	Q6N089_HUMAN	0.57969
2:13	0.973497	EEQYNSTYRVV	0.7218	Q6N089_HUMAN	0.57969
3:1	0.450261	TTEENPIFVVS	0.376	TBK1_HUMAN	6.58705
3:2	0	KVFNNISFLRP	0.7385	TBK1_HUMAN	6.58705
3:3	0.505069	LSLTNQCFDIE	0.6826	TBK1_HUMAN	6.58705
3:4	0.121551	NFFLNSIYKSR	0.6554	ROS1_HUMAN	6.51962
3:5	2.58423	VGLANACYAIH	0.5308	ROS1_HUMAN	6.51962
3:6	96.3106	SGVINESFEGE	1.2859	ROS1_HUMAN	6.51962
3:7	8.322358	DIYKNDYYRKR	0.9268	ROS1_HUMAN	6.51962
3:8	1.380868	SSEENVKYSSS	0.3072	CHK1_HUMAN	2.51454
3:9	2.284798	RREGNIQYLFL	0.3055	CHK1_HUMAN	2.51454
3:10	9.817924	HIQSNLDFSPV	0.3631	CHK1_HUMAN	2.51454
3:11	5.792815	FAIFNLVYWAT	0.9211	Q53F42_HUMAN	0.51696
3:12	2.102681	GKKGNLVYIID	0.4311	KC1D_HUMAN	0.80639
3:13	2.722263	LKPENLLFKDN	0.7907	MAPK5_HUMAN	2.84084

4:1	0.189837	ADGTNTGFPRY	0.8664	RET_HUMAN	2.57367
4:2	0.633538	TWIENKLYGMS	0.8189	RET_HUMAN	2.57367
4:3	2.101196	ITPHNKVYCCD	0.2088	M3K5_HUMAN	0.56847
4:4	1.238038	MCTGNYTFVPY	0.6921	M3K5_HUMAN	0.56847
4:5	1.732325	QNLHNSYYSVP	0.4253	CCNG2_HUMAN	0.32324
4:6	5.498213	QVARNGSFTSI	0.5362	M3K2_HUMAN	13.9926
4:7	1.611094	EDLDNTVFGAE	0.7767	M3K2_HUMAN	13.9926
4:8	0.211608	SDYDNPIFEKF	1.3775	M3K2_HUMAN	13.9926
4:9	2.471434	RKLDNGGYYIT	1.4048	FYN_HUMAN	12.15338
4:10	100.3686	QGFRNAAFEA	1.6076	CJ091_HUMAN	18.67084
4:11	75.73975	FPCGNVNYGYQ	0.3971	ALK_HUMAN	10.82204
4:12	6.301939	MTDYNPNYCFA	0.5677	ALK_HUMAN	10.82204
4:13	88.21687	GGHVNMAFSQS	1.1598	ALK_HUMAN	10.82204
5:1	0	TSLWNPTYGSW	0.6734	ALK_HUMAN	10.82204
5:2	2.846114	DYNPNYCFAGK	0.4522	ALK_HUMAN	10.82204
5:3	51.76767	CGNVNYGYQQQ	1.247	ALK_HUMAN	10.82204
5:4	4.155225	APGSNEQYNVG	0.4221	TNIK_HUMAN	0.78304
5:5	1.211726	NELLNDEFFTS	0.469	PLK1_HUMAN	2.38864
5:6	1.545762	PRGSNPAFLYK	0.9716	BTK_HMUAN	6.83578
5:7	1.268331	VKPENLLYTSK	0.489	MAPK3_HUMAN	3.20892
5:8	1.063448	ITPANVVFLYM	0.428	CDK5_HUMAN	8.61912
5:9	2.984181	KVQPNSSYQNN	0.4397	CDK5_HUMAN	8.61912
5:10	7.809663	TNAENKLYLAE	0.3826	BMX_HUMAN	3.80765
5:11	6.424343	YLAENYCFDSI	0.7079	BMX_HUMAN	3.80765
5:12	2.858528	LTKTNLSYYEY	0.7545	BMX_HUMAN	3.80765
5:13	5.01853	QPNFNMQYIPR	0.2998	BMX_HUMAN	3.80765
6:1	0	YVHVNATYVNV	0.9134	MET_HUMAN	3.19672
6:2	1.371775	LDCSNCPYHIR	0.2644	ABEC4_HUMAN	1.75749
6:3	3.182088	SYPLNNAFPGQ	0.912	ABEC4_HUMAN	1.75749
6:4	1.262137	NPVWNETFVFN	0.9784	KPCG_HUMAN	6.35338
6:5	0.621182	IMEHNVSYPKS	0.3766	KPCA_HUMAN	4.18619
6:6	1.449479	NPQWNESFTFK	0.7417	KPCA_HUMAN	4.18619
6:7	1.810219	FSYVNPQFVHP	1.0977	KPCA_HUMAN	4.18619
6:8	1.524729	IMEHNVAYPKS	0.8052	KPCB_HUMAN	9.92171
6:9	1.722045	NPEWNETFRFQ	0.7249	KPCB_HUMAN	9.92171
6:10	1.856262	HEVKNHKFTAR	0.5356	KPCB_HUMAN	9.92171
6:11	1.263951	FSYTNPEFVIN	0.8792	KPCB_HUMAN	9.92171
6:12	0.818664	EDEANQPFCAV	0.2902	KPCD_HUMAN	4.36358
6:13	1.295992	FSFVNPKFEHL	1.1675	KPCD_HUMAN	4.36358
7:1	0.286851	HYIKNHEFIAT	0.6509	KPCD_HUMAN	4.36358
7:2	1.290646	CNINNFIFHKV	0.7879	KPCD_HUMAN	4.36358

7:3	1.38879	EHVNNTDFKQL	0.4497	CSK21_HUMAN	7.71578
7:4	1.206525	NMFRNFSFMNP	0.955	KPCT_HUMAN	4.30217
7:5	1.366103	QTKENLFFVME	0.572	KPCT_HUMAN	4.30217
7:6	2.788234	FKKTNQFFAIK	0.4389	KPCT_HUMAN	4.30217
7:7	3.410254	RMLMNARYFLE	0.6304	KPCT_HUMAN	4.30217
7:8	2.4609	IRMDNPFYPRW	0.6844	KPCT_HUMAN	4.30217
7:9	1.848984	IREMNPNYTEF	0.393	GSK3A_HUMAN	1.98293
7:10	1.857878	RGEPNVSYICS	0.3656	GSK3A_HUMAN	1.98293
7:11	1.047852	KVIGNGSFGVV	0.9251	GSK3A_HUMAN	1.98293
7:12	1.142243	PTAENPEYLGL	0.3438	ERBB2_HUMAN	3.69018
7:13	10.2265	PAFDNLYYWDQ	1.3717	ERBB2_HUMAN	3.69018
8:1	1.00998	GAVENPEYLTP	0.5545	ERBB2_HUMAN	3.69018
8:2	2.373537	HKATNMEYAVK	0.3672	KS6A1_HUMAN	4.7655
8:3	2.281122	LADVNHPFVVK	0.5991	KS6A1_HUMAN	4.7655
8:4	4.389185	TIDWNKLYRRE	0.3572	KS6A1_HUMAN	4.7655
8:5	5.203229	LHGKNLVFSDG	0.7122	KS6A1_HUMAN	4.7655
8:6	13.5614	LKPSNILYVDE	0.5877	KS6A1_HUMAN	4.7655
8:7	14.47223	VHTPNRTYYLM	0.5936	PDPK1_HUMAN	6.01752
8:8	12.225	ELFENKPYTAT	0.7671	IKKA_HUMAN	5.98267
8:9	4.30782	HSEKNVVYRDL	0.6092	AKT1_HUMAN	6.68603
8:10	1.901632	RDQQNLPYGVT	0.5234	MARK2_HUMAN	1.84412
8:11	1.354292	ANCTNELYMMM	0.518	FGFR2_HUMAN	12.61931
8:12	1.173946	RDINNIDYYKK	0.6759	FGFR2_HUMAN	12.61931
8:13	1.102273	SVKTNPAFLYK	1.0612	FGFR2_HUMAN	12.61931
9:1	0.72588	TLTTNEEYLDL	0.7264	FGFR2_HUMAN	12.61931
9:2	33.68871	SQAINPGFLDS	1.3673	FGR_HUMAN	8.90499
9:3	6.164813	DVCENCHYPIV	0.2215	LCK_HUMAN	16.69021
9:4	96.61679	RNLDNGGFYIS	1.4048	LCK_HUMAN	16.69021
9:5	24.6589	GGYINASFIKI	1.3359	KC1E_HUMAN	1.34505
9:6	100.1917	KGIENKAFDRN	1.5721	JIP4_HUMAN	#N/A
9:7	101.9599	TGFSNMSFEDF	1.2611	Q9U485_DROME (NAK)	#N/A
9:8	103.6306	TGFSNAAFEDF	1.801	NAK Mutant (NAK-AA)	#N/A

S5 OPAL Scoring Matrix of NUMB-PTB

Pos/AA	-5	-4	-2	-1	1	2
K	0.0607	0.09	0.0705	0.0687	0.0752	0.0567
R	0.0607	0.09	0.0705	0.0687	0.0752	0.0567
H	0.0768	0.0129	0.0153	0.0185	0.0675	0.0417
D	0.0017	0.3445	0.0484	0.0149	0.2874	0.3109
E	0	0.1098	0.1888	0.0262	0.2983	0.1158
N	0.0031	0.0218	0.0255	0.0757	0.0878	0.0349
Q	0.0607	0.0306	0.0801	0.021	0.0494	0.0318
S	0.0266	0.0292	0.0363	0.1992	0.0351	0.0292
T	0.0561	0.09	0.0563	0.2231	0.0552	0.0344
A	0.0287	0.0355	0.1778	0.6278	0.0924	0.0567
G	0.0319	0.0335	0.1339	0.5881	0.0752	0.0447
P	0.0567	0.0592	0.08	0.1179	0.0434	0.0476
I	0.4199	0.3052	0.0494	0.0438	0.0428	0.0498
L	0.1753	0.1139	0.1756	0.0625	0.0544	0.0715
V	0.2497	0.3218	0.0644	0.0687	0.079	0.0634
M	0.0939	0.0663	0.0665	0.0587	0.0957	0.081
F	0.357	0.1188	0.1782	0.0659	0.0708	0.1472
Y	0.5542	0.2631	0.2528	0.1023	0.1132	0.1483
W	0.1296	0.0906	0.0705	0.0863	0.0814	0.1071

S6 Prediction of NUMB-PTB binding partners

Rank	Peptide Seq	Uniprot Name	Peptide Position	SMAL
1	IDNDGFED	ITA4_HUMAN	Pos (384-391)	2.0101
2	YKNPAFDD	CT075_HUMAN	Pos (724-731)	1.9503
3	YVNPAFER	PDE8B_HUMAN	Pos (293-300)	1.9388
4	ILNGAFED	S3TC2_HUMAN	Pos (557-564)	1.9047
5	IDNTAFMD	STX3_HUMAN	Pos (26-33)	1.8551
6	YENGAFDE	GP133_HUMAN	Pos (220-227)	1.8289
7	TVNYGFED	OAS3_HUMAN	Pos (270-277)	1.828
8	YGNDGFDD	HNRH3_HUMAN	Pos (161-168)	1.8225
9	YYNEAFSF	SYT5_HUMAN	Pos (305-312)	1.8162
10	YVNAGYAL	AXN1_HUMAN	Pos (305-312)	1.8058
11	ITNVAFDD	MFAP3_HUMAN	Pos (111-118)	1.8004
12	YYNEAYGR	TBB1_HUMAN	Pos (50-57)	1.7658
13	HDNFGFDD	HAX1_HUMAN	Pos (76-83)	1.7859
14	IDNPGYEP	WDR1_HUMAN	Pos (440-447)	1.7784
15	YKNPGYYD	ITLN1_HUMAN	Pos (134-141)	1.7364
		ITLN2_HUMAN	Pos (146-153)	
16	YDNISYED	PKHG1_HUMAN	Pos (661-668)	1.7565
17	YVNGGFAV	EAA3_HUMAN	Pos (503-510)	1.7538
18	LVNEAFDF	RORB_HUMAN	Pos (338-345)	1.7483
19	VYNDGYDD	DYR1A_HUMAN	Pos (135-142)	1.7476
20	FINQAFDM	RIPK5_HUMAN	Pos (373-380)	1.7385
21	YVNWAFSE	TRPM6_HUMAN	Pos (1452-1459)	1.7252
22	LVNEGFEY	HIRA_HUMAN	Pos (946-953)	1.7206
23	YFNYGFNE	FIP1_HUMAN	Pos (175-182)	1.7175
24	YVNSAFTE	TULP4_HUMAN	Pos (1084-1091)	1.7111
25	YPNYAFRE	MMP8_HUMAN	Pos (236-243)	1.685
26	YINGGFTV	SMUF1_HUMAN	Pos (505-512)	1.7
27	IYNQGFEI	CATC_HUMAN	Pos (87-94)	1.6993
28	YKNLAFYW	RISC_HUMAN	Pos (418-425)	1.6679
29	FDNVGYEG	ATP7B_HUMAN	Pos (39-46)	1.697
30	IVNFAFAI	TRPV5_HUMAN	Pos (557-564)	1.6899
31	YNNPAYAD	CV013_HUMAN	Pos (202-209)	1.6871
32	VDNSAFYD	TM179_HUMAN	Pos (161-168)	1.6824
33	YINTAFLK	AOC2_HUMAN	Pos (488-495)	1.6546
34	FDNFAFAM	CAC1C_HUMAN	Pos (346-353)	1.6809
		CAC1D_HUMAN	Pos (347-354)	
35	ITNLGFRD	FAM3C_HUMAN	Pos (173-180)	1.6597
36	YDNTGFVA	LRRC7_HUMAN	Pos (774-781)	1.6788
37	ICNAGFEL	FBN2_HUMAN	Pos (600-607)	1.5556

38	IINAAYKP	SMBT2_HUMAN	Pos (571-578)	1.6535
39	IVNAKFED	FCERA_HUMAN	Pos (80-87)	1.5974
40	YYNQAYFL	ECHP_HUMAN	Pos (487-494)	1.6675
41	YLNEAFSF	CI126_HUMAN	Pos (146-153)	1.667
42	YVNDAFGT	PGK1_HUMAN	Pos (160-167)	1.6618
		PGK2_HUMAN	Pos (160-167)	
43	ICNEGYEG	DNER_HUMAN	Pos (122-129)	1.5398
44	YDNIGFVG	TAB2_HUMAN	Pos (632-639)	1.6599
45	IDNCAFRL	PHKG2_HUMAN	Pos (350-357)	1.5389
46	YGNYAFVW	GRID2_HUMAN	Pos (734-741)	1.6544
47	LVNLAFSD	NK3R_HUMAN	Pos (124-131)	1.6465
48	FVNAAFGK	NET1_HUMAN	Pos (55-62)	1.6163
49	YINASFMD	PTN9_HUMAN	Pos (357-364)	1.643
50	LDNPAFEE	LNX2_HUMAN	Pos (206-213)	1.6417
51	FRNAAFEA	CJ091_HUMAN	Pos (138-145)	1.6076
52	YVNSGFNP	ADRB2_HUMAN	Pos (316-323)	1.6358
53	FVNAAFNV	LAMC1_HUMAN	Pos (54-61)	1.6356
54	FWNAAFFD	K0513_HUMAN	Pos (306-313)	1.6349
55	KINAAFGD	XKR5_HUMAN	Pos (419-426)	1.5576
56	YNNHAFEE	SP17_HUMAN	Pos (69-76)	1.6332
57	IVNEGYVN	SNG3_HUMAN	Pos (43-50)	1.6325
58	ICNGGYEL	FBN3_HUMAN	Pos (1429-1436)	1.5117
59	KVNKAYEF	DNJCD_HUMAN	Pos (1345-1352)	1.5263
60	YANPAFET	PDE8A_HUMAN	Pos (239-246)	1.6302
61	KVNEAFEA	MYOG_HUMAN	Pos (96-103)	1.5541
62	IENKAFDR	JIP4_HUMAN	Pos (370-377)	1.5721
63	QINYAYFD	ERRFI_HUMAN	Pos (206-213)	1.6282
64	LVNLAYES	ASAHL_HUMAN	Pos (115-122)	1.628
65	LDNRAYEM	DPP6_HUMAN	Pos (762-769)	1.5974
66	IANLGFKD	GRIA4_HUMAN	Pos (239-246)	1.6052
67	YYNDAYGA	RBM4B_HUMAN	Pos (210-217)	1.6254
68	YFNFAFKG	KIT_HUMAN	Pos (503-510)	1.5989
69	YKNFAFTM	AT8B2_HUMAN	Pos (896-903)	1.5864
70	YSNMGFHD	Z3H7A_HUMAN	Pos (99-106)	1.6164
71	VSNKAYED	KGP1A_HUMAN	Pos (330-337)	1.5864
		KGP1B_HUMAN	Pos (346-353)	
72	VVNLAYAY	GRK5_HUMAN	Pos (246-253)	1.6156
73	KINEAFEA	MYF6_HUMAN	Pos (108-115)	1.5375
74	IDNDAFRL	OPT_HUMAN	Pos (192-199)	1.5873
75	IVNEGYLN	SNG1_HUMAN	Pos (43-50)	1.6079
76	KVNEAFET	MYOD1_HUMAN	Pos (124-131)	1.5318

77	YKNFAFTL	AT8B1_HUMAN	Pos (956-963)	1.5769
		AT8B4_HUMAN	Pos (878-885)	
78	YINSYFED	RPO3G_HUMAN	Pos (199-206)	1.6072
79	YVNASFID	PTPRA_HUMAN	Pos (588-595)	1.6067
80	IANLAFSD	NPY4R_HUMAN	Pos (80-87)	1.6048
81	IYNWAYGF	PERP_HUMAN	Pos (148-155)	1.6037
82	VRNYGFWD	HPLN1_HUMAN	Pos (235-242)	1.5729
83	IGNLAFSD	NPY5R_HUMAN	Pos (90-97)	1.6028
84	YYNNAYGA	RBM4_HUMAN	Pos (210-217)	1.6025
85	INNAAFEA	PRPS2_HUMAN	Pos (260-267)	1.6023
86	YINASFLD	PTPRF_HUMAN	Pos (1684-1691)	1.6017
87	YENLAFQK	TULP2_HUMAN	Pos (205-212)	1.5735
88	YINFSYEV	AMELX_HUMAN	Pos (28-35)	1.5985
89	PDNEAYEM	SYUA_HUMAN	Pos (120-127)	1.5971
90	PYNYGFDY	ARSF_HUMAN	Pos (154-161)	1.5964
91	IMNRAFFD	KIF3B_HUMAN	Pos (599-606)	1.5662
92	IYNAAYSL	SCNNG_HUMAN	Pos (384-391)	1.5952
93	IVNKAFGI	JIP3_HUMAN	Pos (374-381)	1.565
94	KINAGFGD	RPB1_HUMAN	Pos (1234-1241)	1.5179
95	FVNAAFGQ	PSMD2_HUMAN	Pos (380-387)	1.5914
96	CDNAGFDA	TCPH_HUMAN	Pos (450-457)	1.4545
97	YENGAFQE	DTYMK_HUMAN	Pos (151-158)	1.5909
98	YINASFID	PTPRD_HUMAN	Pos (1699-1706)	1.5901
		PTPRE_HUMAN	Pos (481-488)	
		PTPRS_HUMAN	Pos (1735-1742)	
99	YINASYID	CD45_HUMAN	Pos (705-712)	1.5901
		PTPRE_HUMAN	Pos (189-196)	
100	FTNAAFDP	LYAM3_HUMAN	Pos (821-828)	1.5876
101	YRNGGYIF	PK2L2_HUMAN	Pos (195-202)	1.5562
102	IDNMAFTL	FANCB_HUMAN	Pos (756-763)	1.5854
103	IANLGFTD	GRIA2_HUMAN	Pos (237-244)	1.5852
104	HDNCAYDY	TLL1_HUMAN	Pos (512-519)	1.4848
105	FDNFGFSM	CAC1S_HUMAN	Pos (275-282)	1.5839
106	LVNRAYID	SC15B_HUMAN	Pos (364-371)	1.5491
107	KINEGFDL	ADH1A_HUMAN	Pos (354-361)	1.5017
		ADH1B_HUMAN	Pos (354-361)	
		ADH1G_HUMAN	Pos (354-361)	
108	FDNVGFGY	SCN1A_HUMAN	Pos (1415-1422)	1.5775
109	FDNFGFGT	GRIPE_HUMAN	Pos (924-931)	1.5774
110	YRNLAFLG	ZN254_HUMAN	Pos (46-53)	1.5467
		ZN539_HUMAN	Pos (46-53)	

		ZNF91_HUMAN	Pos (46-53)	
111	YGNFGYGF	AL2SA_HUMAN	Pos (241-248)	1.5764
112	YPNLAFCV	ENTK_HUMAN	Pos (546-553)	1.4802
113	YINGTFDL	UBP47_HUMAN	Pos (83-90)	1.5753
114	IINATYEF	MCAR1_HUMAN	Pos (285-292)	1.5715
115	YINFSYEN	AMELY_HUMAN	Pos (28-35)	1.57
116	FDNQAYPE	PAR15_HUMAN	Pos (642-649)	1.5686
117	FGNFAFDK	PAPD1_HUMAN	Pos (448-455)	1.5406
118	VLNHGYDD	DYR1B_HUMAN	Pos (87-94)	1.5653
119	VDNAGFLY	F109A_HUMAN	Pos (18-25)	1.5628
120	FVNYTFKD	KBTB4_HUMAN	Pos (16-23)	1.5408
121	IINDAFNL	AMPN_HUMAN	Pos (660-667)	1.5606
122	YHNGAYYE	FCGBP_HUMAN	Pos (2789-2796)	1.5578
123	YGNYGYAN	BET1_HUMAN	Pos (15-22)	1.5559
124	YINYIFND	HERC5_HUMAN	Pos (868-875)	1.5547
125	YTNYTFDF	CNOT6_HUMAN	Pos (479-486)	1.5547
126	IRNVAYDT	PLOD3_HUMAN	Pos (240-247)	1.5239
127	YENWAYGE	MRC1_HUMAN	Pos (736-743)	1.5533
128	IENPGFEA	GI24_HUMAN	Pos (240-247)	1.5528
129	FINCAYVK	YLA1_HUMAN	Pos (171-178)	1.4257
130	FRNRGFGD	ZIC1_HUMAN	Pos (102-109)	1.4917
131	PVNLAFDM	HUNK_HUMAN	Pos (699-706)	1.5503
132	LTNYAYDE	TAF1A_HUMAN	Pos (260-267)	1.5491
133	YGNYGYNS	HNRPD_HUMAN	Pos (302-309)	1.5456
134	PVNEGFVD	STAB1_HUMAN	Pos (2371-2378)	1.5453
135	IYNPAFFK	SIA4B_HUMAN	Pos (256-263)	1.5183
136	IVNTAFTV	GTR1_HUMAN	Pos (315-322)	1.5444
137	FDNKGYFC	CRDL1_HUMAN	Pos (329-336)	1.4309
138	KFNYGFEY	DYHC_HUMAN	Pos (1865-1872)	1.467
139	YCNKGFKK	ZN236_HUMAN	Pos (971-978)	1.3447
140	QDNFGYEL	SPTN4_HUMAN	Pos (454-461)	1.5413
141	VVNYPFDD	CBPD_HUMAN	Pos (676-683)	1.5405
142	VKNLGYRD	ARAF_HUMAN	Pos (291-298)	1.4895
143	YLNLGYVS	S3TC2_HUMAN	Pos (125-132)	1.54
144	FDNIGYAW	CAC1G_HUMAN	Pos (337-344)	1.5385
		CAC1H_HUMAN	Pos (361-368)	
		CAC1I_HUMAN	Pos (340-347)	
145	YWNQGYGY	ROAA_HUMAN	Pos (274-281)	1.5365
146	VVNEGFFE	GAPR1_HUMAN	Pos (138-145)	1.535
147	IVNNAYGV	UGTAP_HUMAN	Pos (250-257)	1.5336
148	ITNTGFEM	PZP_HUMAN	Pos (307-314)	1.5336

149	SVNAAFEM	ECHM_HUMAN	Pos (243-250)	1.5333
150	IVNSAFPR	PTPRR_HUMAN	Pos (80-87)	1.5059
151	IVNNAFKH	ACO11_HUMAN	Pos (287-294)	1.5119
152	LQNRAFDD	ASXL1_HUMAN	Pos (857-864)	1.5025
153	KENDAFED	OPTN_HUMAN	Pos (501-508)	1.4559
154	FCNPAFEP	MFRP_HUMAN	Pos (21-28)	1.4107
155	QDNFGFDL	SPTB2_HUMAN	Pos (447-454)	1.5304
156	QDNFGYDL	SPTB1_HUMAN	Pos (446-453)	1.5304
157	YENYSFEE	MYCB2_HUMAN	Pos (2341-2348)	1.5301
158	ILNEAYRK	PACA_HUMAN	Pos (87-94)	1.4823
159	EDNYGYDA	NUP88_HUMAN	Pos (304-311)	1.5295
160	YWNLGFQL	SUIS_HUMAN	Pos (349-356)	1.5294
161	YCNKAFTL	ZBT38_HUMAN	Pos (1070-1077)	1.3792
162	YTNGAYGP	BAG4_HUMAN	Pos (131-138)	1.5287
163	FYNEAFSA	GLT11_HUMAN	Pos (159-166)	1.5285
164	VINAAFML	AP180_HUMAN	Pos (192-199)	1.5277
		PICAL_HUMAN	Pos (192-199)	
165	YNNYGYGN	HNRH3_HUMAN	Pos (156-163)	1.527
166	ICNSGYEV	FBN1_HUMAN	Pos (749-756)	1.406
167	IVNSGFGK	MCM3A_HUMAN	Pos (134-141)	1.498
168	YYNVIYDD	PIWL4_HUMAN	Pos (784-791)	1.5238
169	YCNDGFCE	KCNH2_HUMAN	Pos (43-50)	1.3065
		KCNH6_HUMAN	Pos (43-50)	
		KCNH7_HUMAN	Pos (43-50)	
170	VWNEGFEW	DYSF_HUMAN	Pos (45-52)	1.5226
171	YGNAGYRL	BMF_HUMAN	Pos (97-104)	1.5003
172	YFNQAFHL	ZSWM1_HUMAN	Pos (385-392)	1.5199
173	KVNEAYRF	PGCB_HUMAN	Pos (99-106)	1.4215
174	LDNAAFPF	TFR1_HUMAN	Pos (537-544)	1.516
175	YLNDAFLE	LAP2_HUMAN	Pos (144-151)	1.5145
		LRRC7_HUMAN	Pos (144-151)	
176	LDNSGFEL	CF211_HUMAN	Pos (252-259)	1.514
177	YLNKAFHI	K0056_HUMAN	Pos (684-691)	1.4837
178	FDNQAFLP	MUCDL_HUMAN	Pos (713-720)	1.5114
179	VVNAAYFV	SPB5_HUMAN	Pos (161-168)	1.5113
180	RDNPAYDT	APS_HUMAN	Pos (47-54)	1.4348
181	LDNDGFPD	ITA9_HUMAN	Pos (374-381)	1.5106
182	LVNLAFAE	NK1R_HUMAN	Pos (71-78)	1.5087
183	KENEAFEF	MPIP1_HUMAN	Pos (141-148)	1.4326
184	YSNLGFYP	NR1H4_HUMAN	Pos (77-84)	1.5079
185	GYNFGFEY	M3K4_HUMAN	Pos (1039-1046)	1.5079

186	FPNGAFDH	GLIS1_HUMAN	Pos (601-608)	1.507
187	LSNVAYED	AP4E1_HUMAN	Pos (821-828)	1.5059
188	TVNFAYDT	MYOC_HUMAN	Pos (448-455)	1.5057
189	VCNAGFEL	FBN3_HUMAN	Pos (515-522)	1.3854
190	ICNYGFVG	SUSD1_HUMAN	Pos (58-65)	1.3845
191	RDNAAYFF	SIGL6_HUMAN	Pos (103-110)	1.4288
192	IINHGYAR	CHD4_HUMAN	Pos (1741-1748)	1.4776
193	RDNPGFEA	PKD2_HUMAN	Pos (89-96)	1.4283
194	VDNYMFDD	K2022_HUMAN	Pos (291-298)	1.504
195	FINDAYQE	VP13A_HUMAN	Pos (1274-1281)	1.5036
196	YINANYVD	PTPRG_HUMAN	Pos (906-913)	1.5028
		PTPRS_HUMAN	Pos (1446-1453)	
		PTPRZ_HUMAN	Pos (1781-1788)	
197	ICNEGYYL	MCP_HUMAN	Pos (126-133)	1.3815
198	IKNYGFVH	RBM4B_HUMAN	Pos (34-41)	1.4715
		RBM4_HUMAN	Pos (34-41)	
199	YPNQAFKR	PA24A_HUMAN	Pos (685-692)	1.4532
200	LVNYKYED	DOCK9_HUMAN	Pos (123-130)	1.4278
201	IQNAAFMY	KPCD1_HUMAN	Pos (793-800)	1.5001
		KPCD2_HUMAN	Pos (761-768)	
		KPCD3_HUMAN	Pos (786-793)	
202	KVNQAFET	MYF5_HUMAN	Pos (98-105)	1.4231
203	ISNGGFES	KCC2A_HUMAN	Pos (361-368)	1.4986
204	YLNNAYAR	CLIC6_HUMAN	Pos (672-679)	1.4705
205	TINAGFEL	USP9X_HUMAN	Pos (1490-1497)	1.497
		USP9Y_HUMAN	Pos (1499-1506)	
206	IDNSGYVS	PLSI_HUMAN	Pos (25-32)	1.497
207	DVNCAFEF	TR100_HUMAN	Pos (333-340)	1.3968
208	LENYGFDK	PMS1_HUMAN	Pos (44-51)	1.4701
209	YVNFNFTD	CCR9_HUMAN	Pos (28-35)	1.496
210	FLNVAFDG	WDR67_HUMAN	Pos (45-52)	1.4952
211	YENRAYTP	TTC3_HUMAN	Pos (367-374)	1.4651
212	YKNDAYFL	CAH11_HUMAN	Pos (212-219)	1.4627
213	YWNLGFTS	SIA7B_HUMAN	Pos (211-218)	1.4929
214	VVNGAFMV	WDR22_HUMAN	Pos (321-328)	1.4923
215	QLNPAFED	ERCC8_HUMAN	Pos (380-387)	1.4916
216	VCNLGYEA	FBN3_HUMAN	Pos (708-715)	1.3684
217	YVNYSFRV	NRCAM_HUMAN	Pos (714-721)	1.4666
218	KDNTGYDL	D2HDH_HUMAN	Pos (244-251)	1.4085
219	YVNANYID	PTPRD_HUMAN	Pos (1410-1417)	1.4832
220	FTNIAYDL	HS2ST_HUMAN	Pos (89-96)	1.4831

221	RVNSAYQD	PGCA_HUMAN	Pos (95-102)	1.4069
222	YLNGGFII	ATS12_HUMAN	Pos (752-759)	1.4827
223	YLNIAFTM	CAC1E_HUMAN	Pos (1511-1518)	1.4815
224	VFNWAFEE	SIG10_HUMAN	Pos (165-172)	1.4809
225	YKNVAYVN	AT10D_HUMAN	Pos (1119-1126)	1.4503
226	KDNEGFGF	SHAN2_HUMAN	Pos (43-50)	1.4045
227	VDNGGFVC	FATH_HUMAN	Pos (4071-4078)	1.3952
228	QINPAFAD	OGT1_HUMAN	Pos (423-430)	1.477
229	YKNDAYLL	CAH10_HUMAN	Pos (210-217)	1.4463
230	AINKAYEE	RBL2_HUMAN	Pos (333-340)	1.4463
231	QVNKAFLD	PERL_HUMAN	Pos (44-51)	1.4461
232	FDNQGYFN	WDR22_HUMAN	Pos (271-278)	1.4754
233	DVNLAYEI	T2EA_HUMAN	Pos (196-203)	1.475
234	LLNLAYEK	TM16C_HUMAN	Pos (609-616)	1.4476
235	FVNVGYFL	TSN1_HUMAN	Pos (53-60)	1.4736
236	GMNAGFED	KMO_HUMAN	Pos (316-323)	1.4733
237	FCNPGFES	EMR1_HUMAN	Pos (198-205)	1.3526
238	FHNKGYEE	SYT14_HUMAN	Pos (218-225)	1.4426
239	YPNCAYRT	RNAS1_HUMAN	Pos (120-127)	1.3508
240	IFNFAFWG	FPRL2_HUMAN	Pos (177-184)	1.4708
241	ILNAAYHV	RIPK5_HUMAN	Pos (520-527)	1.4703
242	VWNMAFDF	REV1_HUMAN	Pos (1224-1231)	1.4692
243	YLNTAYGH	TM6S1_HUMAN	Pos (102-109)	1.4691
244	YTNWAYVP	ENK1_HUMAN	Pos (100-107)	1.4691
		ENK4_HUMAN	Pos (101-108)	
		ENK6_HUMAN	Pos (101-108)	
245	VVNAAYSA	ALG12_HUMAN	Pos (353-360)	1.4689
246	YWNRGFVA	ABCA6_HUMAN	Pos (170-177)	1.4391
247	IPNGAFVY	NGL1_HUMAN	Pos (139-146)	1.4681
248	IYNKGFLH	ELTD1_HUMAN	Pos (523-530)	1.4377
249	LTNAGYDY	RIOK2_HUMAN	Pos (75-82)	1.4669
250	YINANYID	PTPRF_HUMAN	Pos (1395-1402)	1.4666
		PTPRK_HUMAN	Pos (940-947)	
		PTPRT_HUMAN	Pos (964-971)	
		PTPRU_HUMAN	Pos (927-934)	
251	VVNFPYDD	SORCN_HUMAN	Pos (183-190)	1.4659
252	IPNYGYRI	CSMD1_HUMAN	Pos (1223-1230)	1.445
253	YKNRGFLP	FGF19_HUMAN	Pos (154-161)	1.4048
254	IVNLSFSD	NPY1R_HUMAN	Pos (79-86)	1.4625
255	IRNRAFYM	PDLI7_HUMAN	Pos (374-381)	1.4024
256	LVNYAYTS	KLH23_HUMAN	Pos (91-98)	1.4621

257	FVNTGFTK	DHB13_HUMAN	Pos (220-227)	1.4351
258	TDNVAFDM	GOLI_HUMAN	Pos (321-328)	1.4612
259	RVNAAFYR	LOXL3_HUMAN	Pos (224-231)	1.358
260	HDNIAFQD	LRMP_HUMAN	Pos (143-150)	1.4588
261	LTNLAFVD	OR5MB_HUMAN	Pos (63-70)	1.4586
262	EVNLAYEN	DYHC_HUMAN	Pos (522-529)	1.4584
263	YDNCSFYF	TPTE_HUMAN	Pos (489-496)	1.3583
264	IRNAAFLA	VASH2_HUMAN	Pos (84-91)	1.4266
265	QPNRAFED	GPR98_HUMAN	Pos (5397-5404)	1.4274
266	LENAAYLD	PC11Y_HUMAN	Pos (450-457)	1.456
267	YVNFKFYF	ACA_HUMAN	Pos (64-71)	1.3833
268	ADNAGYDS	TCPB_HUMAN	Pos (450-457)	1.4557
269	YGNEGYIP	TEC_HUMAN	Pos (223-230)	1.455
270	LVNAAFSE	LGI3_HUMAN	Pos (71-78)	1.4536
271	LLNYGYDT	ASB14_HUMAN	Pos (155-162)	1.4519
272	YQNGAFGS	SOX14_HUMAN	Pos (172-179)	1.4509
273	YINEQYEK	SEPT3_HUMAN	Pos (132-139)	1.4242
274	FVNTGFIK	DHRS8_HUMAN	Pos (220-227)	1.4227
275	IFNKGFGF	VDAC2_HUMAN	Pos (43-50)	1.4197
276	LVNAAFYQ	ST18_HUMAN	Pos (651-658)	1.4477
277	IANYAYGN	CRAC1_HUMAN	Pos (195-202)	1.4461
278	YLNTGYQR	A2MG_HUMAN	Pos (1007-1014)	1.4186
279	ILNPAFLY	TRRAP_HUMAN	Pos (1776-1783)	1.4443
280	VCNVAFEK	PPR3D_HUMAN	Pos (195-202)	1.2969
281	LDNLGYNL	PLVAP_HUMAN	Pos (250-257)	1.4428
282	LINEAFRI	ECHP_HUMAN	Pos (626-633)	1.4221
283	LCNEGYEL	FBN2_HUMAN	Pos (1999-2006)	1.322
284	FQNMAFQD	SYNE2_HUMAN	Pos (3606-3613)	1.4422
285	SINLGFSD	SPT16_HUMAN	Pos (378-385)	1.4415
286	LVNLAFCG	O4F29_HUMAN	Pos (163-170)	1.3452
		OR4F3_HUMAN	Pos (163-170)	
287	RDNNGYSD	PCLO_HUMAN	Pos (4593-4600)	1.3648
288	IRNVTFED	FGFR2_HUMAN	Pos (329-336)	1.4066
289	LDNFAYLA	GLTL1_HUMAN	Pos (245-252)	1.4369
290	GVNDAFKD	FUBP1_HUMAN	Pos (26-33)	1.416
291	YWNVGFLK	PIGV_HUMAN	Pos (311-318)	1.4084
292	IENQAFYK	FRMD5_HUMAN	Pos (288-295)	1.4075
293	YVNHSFEG	NPCL1_HUMAN	Pos (1333-1340)	1.4335
294	ITNLGYKV	AASS_HUMAN	Pos (50-57)	1.4122
295	YKNPVYED	MYPC1_HUMAN	Pos (1033-1040)	1.4021
296	LGNEAFKD	TDRD7_HUMAN	Pos (167-174)	1.4115

297	FENAAFGR	LAMC3_HUMAN	Pos (39-46)	1.4043
298	LCNKAFRD	ACM1_HUMAN	Pos (420-427)	1.2597
299	LINAAYKP	SMBT1_HUMAN	Pos (545-552)	1.4089
300	KDNPAFAF	SF04_HUMAN	Pos (215-222)	1.3526
301	TVNSAYGD	CD166_HUMAN	Pos (30-37)	1.4281
302	YPNRGFPK	METH_HUMAN	Pos (988-995)	1.3721
303	GINTAFCD	NOTC1_HUMAN	Pos (926-933)	1.3321
304	LDNGAYNA	ARP6_HUMAN	Pos (6-13)	1.426
305	IDNYVYEH	AZIN1_HUMAN	Pos (25-32)	1.4259
306	IINENYDY	NUP98_HUMAN	Pos (1659-1666)	1.4253
307	YANDAFPE	RHOJ_HUMAN	Pos (41-48)	1.4251
		RHOQ_HUMAN	Pos (29-36)	
308	VINGGFGL	HUTU_HUMAN	Pos (600-607)	1.4236
309	FRNLGFGW	CP11A_HUMAN	Pos (450-457)	1.393
310	YWNQGYGN	HNRPD_HUMAN	Pos (297-304)	1.4231
311	IGNAGFNE	DDEF2_HUMAN	Pos (492-499)	1.4229
312	YINGNYID	PTPRM_HUMAN	Pos (953-960)	1.4227
313	ISNGAYRE	NUAK2_HUMAN	Pos (262-269)	1.4018
314	KDNGAYRC	FCRL5_HUMAN	Pos (159-166)	1.2421
315	LENFGFEL	ATBF1_HUMAN	Pos (1949-1956)	1.4212
		ZFHX2_HUMAN	Pos (299-306)	
316	MCNLAFAD	FSHR_HUMAN	Pos (401-408)	1.3006
		TSHR_HUMAN	Pos (453-460)	
317	ILNGAFRS	PHF2_HUMAN	Pos (401-408)	1.3999
318	YSNIAFNL	ESPL1_HUMAN	Pos (730-737)	1.4199
319	LINFAYNG	KLH18_HUMAN	Pos (27-34)	1.419
320	VDNVGFNR	ZBED4_HUMAN	Pos (676-683)	1.3912
321	VGNKAFRD	PKHG4_HUMAN	Pos (1046-1053)	1.3676
322	KINEAFIE	MATR3_HUMAN	Pos (433-440)	1.3411
323	YTNYCYHD	ZSWM6_HUMAN	Pos (68-75)	1.2754
324	IKNGAFVN	ANFY1_HUMAN	Pos (309-316)	1.3855
325	FVNYYFER	PHF14_HUMAN	Pos (619-626)	1.3889
326	IENCGFQY	DNMT2_HUMAN	Pos (137-144)	1.3155
327	WINIGYEG	MARK4_HUMAN	Pos (309-316)	1.4153
328	VVNTAYGR	NLGN2_HUMAN	Pos (43-50)	1.3875
329	LDNLGYRT	ADDA_HUMAN	Pos (384-391)	1.3931
		ADDB_HUMAN	Pos (371-378)	
		ADDG_HUMAN	Pos (378-385)	
330	WDNEAYVH	IL1B_HUMAN	Pos (108-115)	1.4114
331	GINRGFGD	MYEF2_HUMAN	Pos (423-430)	1.3818
332	VRNMGFGD	GRHL3_HUMAN	Pos (592-599)	1.3804

333	LINLAYLL	ERCC5_HUMAN	Pos (851-858)	1.4098
334	VVNMGYSY	UBAP1_HUMAN	Pos (399-406)	1.4095
335	YCNMGFSG	ELTD1_HUMAN	Pos (43-50)	1.2886
336	GVNGGYET	ANR38_HUMAN	Pos (667-674)	1.4084
337	YRNLTFDP	TRI65_HUMAN	Pos (330-337)	1.3779
338	LANLAFWD	GPR37_HUMAN	Pos (300-307)	1.4065
339	VPNKAFEL	AN30A_HUMAN	Pos (477-484)	1.377
340	PFNEAFEE	LCB2_HUMAN	Pos (58-65)	1.4062
341	KDNFGFIE	CSDE1_HUMAN	Pos (527-534)	1.3301
342	YINRNFDR	DEPD5_HUMAN	Pos (321-328)	1.3497
343	RMNEAFGD	TIF1B_HUMAN	Pos (794-801)	1.3297
344	LDNGGFYI	HCK_HUMAN	Pos (202-209)	1.4048
		LCK_HUMAN	Pos (185-192)	
345	LDNGGYYI	FYN_HUMAN	Pos (207-214)	1.4048
		LYN_HUMAN	Pos (187-194)	
		YES_HUMAN	Pos (216-223)	
346	YPNETFEM	ABCAC_HUMAN	Pos (2176-2183)	1.4046
347	VSNFGYDL	PEDF_HUMAN	Pos (58-65)	1.4041
348	YDNPIFEK	M3K2_HUMAN	Pos (250-257)	1.3775
349	GVNGGYES	ANR25_HUMAN	Pos (503-510)	1.4032
350	EINKAFEL	ADHX_HUMAN	Pos (353-360)	1.3733
351	YVNFDFDG	IF36_HUMAN	Pos (300-307)	1.4012
352	YDNFSYRS	CX045_HUMAN	Pos (90-97)	1.3805
353	YINRTFFF	POMT2_HUMAN	Pos (96-103)	1.371
354	VWNEAFRF	PA24D_HUMAN	Pos (71-78)	1.3793
355	RDNAGYCA	BCOR_HUMAN	Pos (1493-1500)	1.2278
356	FRNAGYAV	SYJ2B_HUMAN	Pos (85-92)	1.3687
357	YGNKYYED	NDUAC_HUMAN	Pos (44-51)	1.3697
358	YSNDGFCK	KCNH1_HUMAN	Pos (44-51)	1.2766
		KCNH5_HUMAN	Pos (42-49)	
359	HENAAFKD	TOIP1_HUMAN	Pos (498-505)	1.3783
		TOIP2_HUMAN	Pos (384-391)	
360	LINAGYRA	SARDH_HUMAN	Pos (760-767)	1.3783
361	PNNGGYDD	DYRK2_HUMAN	Pos (125-132)	1.3988
		DYRK3_HUMAN	Pos (185-192)	
362	LSNLAFVD	OR1C1_HUMAN	Pos (63-70)	1.3978
		OR8U1_HUMAN	Pos (63-70)	
363	SINTAYER	MMP14_HUMAN	Pos (367-374)	1.3709
364	YMNDGYAP	DLG1_HUMAN	Pos (409-416)	1.397
365	ICNEGYSG	TENA_HUMAN	Pos (610-617)	1.2766
366	KVNYNFED	OAS2_HUMAN	Pos (609-616)	1.3202

367	FENFKYED	PRAX_HUMAN	Pos (72-79)	1.3229
368	GVNMGFGD	COQ6_HUMAN	Pos (378-385)	1.3944
369	VKNYAYKP	IPO11_HUMAN	Pos (328-335)	1.3431
370	WINAAFRA	COG7_HUMAN	Pos (16-23)	1.3723
371	GINGGYET	ANR15_HUMAN	Pos (998-1005)	1.3918
372	ICNGAFTT	ZN236_HUMAN	Pos (717-724)	1.2712
373	LDNGGFAA	SDK1_HUMAN	Pos (2056-2063)	1.3909
374	VINPAYAT	KRIT1_HUMAN	Pos (190-197)	1.3895
375	KVNLAFKQ	VP26A_HUMAN	Pos (51-58)	1.2929
376	GENYGYRD	CSMD3_HUMAN	Pos (2793-2800)	1.3687
377	YANSAFNP	ADRB1_HUMAN	Pos (367-374)	1.3892
		ADRB3_HUMAN	Pos (336-343)	
378	ITNGAYNS	TBR1_HUMAN	Pos (152-159)	1.3886
379	LDNVAYAR	RAD51_HUMAN	Pos (186-193)	1.3611
380	YCNRDFDD	ZN207_HUMAN	Pos (15-22)	1.2379
381	LVNLGFAT	ASCL1_HUMAN	Pos (133-140)	1.3876
382	YYNESFSF	SYT1_HUMAN	Pos (339-346)	1.3876
383	TINLAFGF	AQP3_HUMAN	Pos (58-65)	1.3871
384	TGNKGFED	MEPE_HUMAN	Pos (121-128)	1.3574
385	ILNLGFLT	S22A3_HUMAN	Pos (160-167)	1.3863
386	PTNYGYND	ATS8_HUMAN	Pos (699-706)	1.3863
387	IENGAFQG	PGS2_HUMAN	Pos (191-198)	1.3855
388	IENAVFED	E2AK3_HUMAN	Pos (1071-1078)	1.3854
389	KMNEGFEW	DUS12_HUMAN	Pos (153-160)	1.3093
390	IPNEAFLN	TLR8_HUMAN	Pos (654-661)	1.385
391	VTNQAFDI	SCN9A_HUMAN	Pos (1512-1519)	1.3848
392	MVNLAFVK	UBP19_HUMAN	Pos (332-339)	1.3548
393	LYNAAFAG	LMBRL_HUMAN	Pos (438-445)	1.3811
394	LANLGFMD	GRIA1_HUMAN	Pos (230-237)	1.3811
395	YINKYYDQ	PLXB1_HUMAN	Pos (2094-2101)	1.3514
396	WDNAGFKG	ADA17_HUMAN	Pos (267-274)	1.3599
397	KDNRAYNC	PCD12_HUMAN	Pos (745-752)	1.1913
398	LANLAFLD	OR4M2_HUMAN	Pos (63-70)	1.3795
399	TINEAFLF	CCNB3_HUMAN	Pos (1002-1009)	1.3795
400	PVNNAFSD	POSTN_HUMAN	Pos (408-415)	1.3778
401	LGNLAFLD	OR4KE_HUMAN	Pos (64-71)	1.3775
		OR4N2_HUMAN	Pos (63-70)	
		OR4N4_HUMAN	Pos (63-70)	
402	ADNNGYVD	ITA4_HUMAN	Pos (446-453)	1.3767
403	FLNGGYSY	ADCY8_HUMAN	Pos (142-149)	1.3763
404	IPNGAFAH	HES1_HUMAN	Pos (233-240)	1.3749

405	FENLGFRP	FAKD2_HUMAN	Pos (426-433)	1.3533
406	VLNDAYEN	SYNE2_HUMAN	Pos (3250-3257)	1.373
407	AINTAFEA	DCR1A_HUMAN	Pos (858-865)	1.373
408	NYNVAFEE	SPEG_HUMAN	Pos (1811-1818)	1.3725
409	IANLAYVT	MOV10_HUMAN	Pos (63-70)	1.3722
410	SINEAYGY	NR1I3_HUMAN	Pos (325-332)	1.3719
411	FDNLYYWD	ERBB2_HUMAN	Pos (1217-1224)	1.3717
412	RVNIAFNY	UAP56_HUMAN	Pos (355-362)	1.2958
413	SVNQGFEA	SMAD2_HUMAN	Pos (397-404)	1.3716
		SMAD3_HUMAN	Pos (355-362)	
414	YSNCGYSV	PTPRT_HUMAN	Pos (38-45)	1.27
415	RINLAFMS	K0652_HUMAN	Pos (190-197)	1.2942
416	YLNSGYIT	NPT4_HUMAN	Pos (291-298)	1.3697
417	FKNKAFAL	IPYR2_HUMAN	Pos (258-265)	1.3092
418	PDNLGYTY	TYRP1_HUMAN	Pos (457-464)	1.3684
419	IFNDAFWL	ATM_HUMAN	Pos (1927-1934)	1.3678
420	AINPGFLD	FGR_HUMAN	Pos (60-67)	1.3673
421	PYNLAYKY	RENR_HUMAN	Pos (292-299)	1.3467
422	IKNMAFQK	AQR_HUMAN	Pos (1104-1111)	1.3103
423	LVNLGFQA	ASCL2_HUMAN	Pos (65-72)	1.3669
424	DINMAFLD	DSG4_HUMAN	Pos (773-780)	1.3665
425	LGNLAFID	OR5V1_HUMAN	Pos (63-70)	1.3659
426	YGNVAFPH	FIP1_HUMAN	Pos (426-433)	1.365
427	ILNVAYGV	DUS19_HUMAN	Pos (95-102)	1.3646
428	FLNFGFAN	PTN22_HUMAN	Pos (783-790)	1.3645
429	YENHGYQP	TMPS2_HUMAN	Pos (15-22)	1.3644
430	FWNLGYWL	ACM3_HUMAN	Pos (525-532)	1.3642
431	DVNWGYEK	SCG1_HUMAN	Pos (579-586)	1.3371
432	LINKGFVE	SYDC_HUMAN	Pos (210-217)	1.3339
433	FTNLGYRA	THOC6_HUMAN	Pos (329-336)	1.3426
434	ISNYGFPS	MMP8_HUMAN	Pos (365-372)	1.3626
435	YINASFIK	PTN13_HUMAN	Pos (2262-2269)	1.3359
436	AINEAYKE	ANXA6_HUMAN	Pos (471-478)	1.3415
437	LSNLAFID	OR8I2_HUMAN	Pos (63-70)	1.3616
438	RANLAYGD	PITM2_HUMAN	Pos (534-541)	1.2857
439	YYNTCFSD	SASH1_HUMAN	Pos (69-76)	1.2196
440	YNNKGYHS	ABCA2_HUMAN	Pos (1732-1739)	1.3313
441	VDNKGFPT	EZRI_HUMAN	Pos (49-56)	1.3306
442	YINASFIM	CD45_HUMAN	Pos (1015-1022)	1.3602
443	YINASYIM	PTPRG_HUMAN	Pos (1205-1212)	1.3602
		PTPRZ_HUMAN	Pos (2076-2083)	

444	LYNIAFMY	NCF2_HUMAN	Pos (124-131)	1.3596
445	CVNEGYAQ	ASCL3_HUMAN	Pos (107-114)	1.2229
446	YDNQEFPD	CP2E1_HUMAN	Pos (398-405)	1.3593
447	LLNSGYEF	ATPG_HUMAN	Pos (173-180)	1.3591
448	KCNMGYEY	CFAH_HUMAN	Pos (236-243)	1.1619
449	KINGAYFC	TIE2_HUMAN	Pos (95-102)	1.1984
450	IDNYSYRG	B3GL1_HUMAN	Pos (210-217)	1.3363
451	LLNNAFEE	MCM3_HUMAN	Pos (65-72)	1.3566
452	IWNTGYLF	CAH5A_HUMAN	Pos (95-102)	1.3565
453	ALNFAFKD	ADT1_HUMAN	Pos (85-92)	1.3347
		ADT2_HUMAN	Pos (85-92)	
		ADT3_HUMAN	Pos (85-92)	
454	FLNGRFED	RHES_HUMAN	Pos (39-46)	1.2827
455	ILNNAFRL	NFKB1_HUMAN	Pos (843-850)	1.3338
456	KDNMAYMC	ASPM_HUMAN	Pos (398-405)	1.1952
457	IGNGAYGV	MK07_HUMAN	Pos (60-67)	1.3537
458	KVNAAFTT	TRPC3_HUMAN	Pos (606-613)	1.2777
459	MVNYNFED	OAS2_HUMAN	Pos (265-272)	1.3534
460	VLNSGFLD	FA62A_HUMAN	Pos (741-748)	1.3533
461	YHNPGYLV	LAT_HUMAN	Pos (156-163)	1.353
462	IVNEIYHD	ATS2_HUMAN	Pos (297-304)	1.3527
		ATS3_HUMAN	Pos (287-294)	
463	KENGAYDA	CSN3_HUMAN	Pos (170-177)	1.2763
464	ICNTGYRL	FHR1_HUMAN	Pos (113-120)	1.211
		FHR2_HUMAN	Pos (113-120)	
465	TDNRAFGY	DAB2_HUMAN	Pos (129-136)	1.3224
466	LANFAYDP	ARMC7_HUMAN	Pos (40-47)	1.3518
467	KDNLGFPV	FACE1_HUMAN	Pos (440-447)	1.2757
468	SCNAAFND	FBX42_HUMAN	Pos (143-150)	1.2309
469	SINDAFEG	PTF1A_HUMAN	Pos (178-185)	1.351
470	IKNHAFRL	GAMT_HUMAN	Pos (152-159)	1.2997
471	YRNITFDE	METH_HUMAN	Pos (164-171)	1.3199
472	YINASYVN	PTN3_HUMAN	Pos (695-702)	1.3503
473	IVNLLYEL	RYR1_HUMAN	Pos (517-524)	1.3496
474	AENEAFDW	TTBK1_HUMAN	Pos (303-310)	1.3496
475	HMNEAFVD	CAN12_HUMAN	Pos (199-206)	1.3496
476	ICNKAYQQ	ZBT40_HUMAN	Pos (1010-1017)	1.1994
477	EKNYAFEI	DPOLA_HUMAN	Pos (436-443)	1.3187
478	KENPGFDF	NUDC2_HUMAN	Pos (129-136)	1.2732
479	ICNPGYKG	TEN4_HUMAN	Pos (647-654)	1.2079
480	VDNHGYNV	SUIS_HUMAN	Pos (108-115)	1.3488

481	ICNPGYMG	NOTC2_HUMAN	Pos (594-601)	1.2284
482	LSNWAYEF	SMCA4_HUMAN	Pos (815-822)	1.3483
483	LLNPAYDV	PO210_HUMAN	Pos (244-251)	1.3478
484	VINGTYDY	ALG3_HUMAN	Pos (81-88)	1.3476
485	YENEMFDY	DOCK1_HUMAN	Pos (1304-1311)	1.3472
486	FPNLAFIR	PRELP_HUMAN	Pos (265-272)	1.3191
487	YINVWYDP	MK08_HUMAN	Pos (320-327)	1.3451
		MK10_HUMAN	Pos (358-365)	
488	PTNEAFEK	BGH3_HUMAN	Pos (280-287)	1.3183
		POSTN_HUMAN	Pos (274-281)	
489	IINDTYEI	CHL1_HUMAN	Pos (955-962)	1.3447
490	YVNLLFWY	AT10B_HUMAN	Pos (1123-1130)	1.3438
		AT10D_HUMAN	Pos (1124-1131)	
491	ICNSGFKR	I15RA_HUMAN	Pos (58-65)	1.1762
492	SYNTAFEL	IL1F9_HUMAN	Pos (158-165)	1.3436
493	YINKKFCF	SYT14_HUMAN	Pos (45-52)	1.1458
494	FINGSFEI	SO1A2_HUMAN	Pos (60-67)	1.3434
495	VVNRLFED	TXND3_HUMAN	Pos (577-584)	1.3137
496	EYNTGYDY	RBM6_HUMAN	Pos (491-498)	1.3432
497	ISNGAYGA	MAST1_HUMAN	Pos (380-387)	1.3427
		MAST2_HUMAN	Pos (518-525)	
		MAST3_HUMAN	Pos (373-380)	
		MAST4_HUMAN	Pos (397-404)	
498	IFNDGYKL	FGL1_HUMAN	Pos (86-93)	1.3219
499	FLNDAYNW	PHTF1_HUMAN	Pos (638-645)	1.342
500	VWNMGFSD	MBTP1_HUMAN	Pos (764-771)	1.3409
501	LANLGFTD	GRIA3_HUMAN	Pos (247-254)	1.3406
502	GVNLGFGF	AQP7_HUMAN	Pos (69-76)	1.3398
503	PLNLAFLD	STAG2_HUMAN	Pos (994-1001)	1.3393
504	INNEAFSG	TDRD3_HUMAN	Pos (461-468)	1.3381
505	VLNLAYNK	TLR5_HUMAN	Pos (316-323)	1.3115
506	LSNAAFES	KI18A_HUMAN	Pos (592-599)	1.3376
507	ANNIAYED	SCG2_HUMAN	Pos (236-243)	1.3369
508	LINDAYKR	AFG32_HUMAN	Pos (704-711)	1.2886
509	LDNPAYVS	LNX1_HUMAN	Pos (179-186)	1.3358
510	DINEAFKE	HTF4_HUMAN	Pos (592-599)	1.3145
		ITF2_HUMAN	Pos (579-586)	
		TFE2_HUMAN	Pos (564-571)	
511	EDNAAYYL	GNAZ_HUMAN	Pos (149-156)	1.3348
512	YYNEYFVF	OTOF_HUMAN	Pos (296-303)	1.3346
513	YDNCPFFF	TPTE2_HUMAN	Pos (471-478)	1.2346

514	ELNFAFEE	NAL14_HUMAN	Pos (263-270)	1.334
515	INNLAFLT	RYR3_HUMAN	Pos (3321-3328)	1.3339
516	YDNLCYLV	PKDRE_HUMAN	Pos (1227-1234)	1.1921
517	FTNQGYFF	IL2RB_HUMAN	Pos (359-366)	1.3332
518	YVNRRFHF	EDG2_HUMAN	Pos (75-82)	1.2299
519	GPNAAYDF	MAGD1_HUMAN	Pos (121-128)	1.3313
520	FKNKAYLL	DICER_HUMAN	Pos (1667-1674)	1.2712
521	ILNVAFTI	CAC1S_HUMAN	Pos (1152-1159)	1.331
522	FDNVSYLD	TOPRS_HUMAN	Pos (110-117)	1.3304
523	IYNFPFDV	5HT3R_HUMAN	Pos (161-168)	1.3299
524	YVNACFFV	SMO_HUMAN	Pos (269-276)	1.188
525	IDNGVFEV	GRP78_HUMAN	Pos (237-244)	1.3287
526	YHNIAFIH	AAKG2_HUMAN	Pos (433-440)	1.3288
527	HCNRAFAD	SLUG_HUMAN	Pos (217-224)	1.1784
528	YCNFVFKD	PLD1_HUMAN	Pos (649-656)	1.1872
529	KLNWAFEM	HPCL4_HUMAN	Pos (99-106)	1.2522
530	RDNLGFSP	KCNKF_HUMAN	Pos (301-308)	1.2516
531	SCNAGYEE	EPHA3_HUMAN	Pos (258-265)	1.2066
532	YINASYIP	PTPRB_HUMAN	Pos (1757-1764)	1.3268
533	IVNEFYEQ	PDE3A_HUMAN	Pos (968-975)	1.3265
		PDE3B_HUMAN	Pos (955-962)	
534	VYNFEFED	CM031_HUMAN	Pos (126-133)	1.3264
535	YINEYYAK	ABCA2_HUMAN	Pos (1966-1973)	1.2996
536	INNTGFAF	FOXJ2_HUMAN	Pos (403-410)	1.3257
537	MYNFAYVK	SCN7A_HUMAN	Pos (1399-1406)	1.2987
538	YSNSGFWN	MRC1_HUMAN	Pos (903-910)	1.3241
539	PYNEGYVY	ETV1_HUMAN	Pos (470-477)	1.324
540	SVNEAFGR	PWP1_HUMAN	Pos (465-472)	1.2969
541	PENFAFDV	LBXCO_HUMAN	Pos (176-183)	1.3233
542	DDNYAFLH	ABC3F_HUMAN	Pos (174-181)	1.3229
543	RINLGFNS	MXRA5_HUMAN	Pos (59-66)	1.2466
544	KDNKGYCA	MUSK_HUMAN	Pos (311-318)	1.1205
545	LDNIAYIM	HM2L1_HUMAN	Pos (591-598)	1.3208
546	YRNQTFEG	DIDO1_HUMAN	Pos (2078-2085)	1.2904
547	QCNEGYEV	FBN1_HUMAN	Pos (1957-1964)	1.1993
548	LENMAYEH	DDX3X_HUMAN	Pos (570-577)	1.3194
		DDX3Y_HUMAN	Pos (569-576)	
549	IINIKFAD	FLNA_HUMAN	Pos (2114-2121)	1.2465
		FLNC_HUMAN	Pos (2109-2116)	
550	LYNLGFGH	ITIH1_HUMAN	Pos (430-437)	1.319
551	DINEAYVE	ATG4B_HUMAN	Pos (234-241)	1.3183

552	LVNMAYTL	ITBA1_HUMAN	Pos (59-66)	1.3181
553	YTNQTFDK	RPC2_HUMAN	Pos (782-789)	1.2915
554	IENIGYVL	TRPC3_HUMAN	Pos (643-650)	1.3177
		TRPC6_HUMAN	Pos (700-707)	
		TRPC7_HUMAN	Pos (645-652)	
555	DYNEAYNY	DNJC7_HUMAN	Pos (43-50)	1.3175
556	LMNIAFND	TBCD2_HUMAN	Pos (851-858)	1.3175
557	FDNRSYDS	ENAH_HUMAN	Pos (524-531)	1.2878
558	FINYTYKR	STK38_HUMAN	Pos (440-447)	1.27
559	ILNVAFPP	NTCP_HUMAN	Pos (269-276)	1.317
560	YINYLFYS	GBRT_HUMAN	Pos (348-355)	1.3171
561	YNNPLYDD	PARM1_HUMAN	Pos (302-309)	1.3168
562	KENKGYFD	ATRX_HUMAN	Pos (1896-1903)	1.2108
563	YSNPGYSS	RFFL_HUMAN	Pos (29-36)	1.3158
564	LDNVGFKP	GAS8_HUMAN	Pos (449-456)	1.2951
565	YDNPSFLR	ZDH17_HUMAN	Pos (286-293)	1.289
566	ILNLSFCD	GPR75_HUMAN	Pos (83-90)	1.2195
567	ILNRGFPI	CENPJ_HUMAN	Pos (30-37)	1.2856
568	IANYSFKD	ITA8_HUMAN	Pos (237-244)	1.2935
569	YINASFIN	PTPRA_HUMAN	Pos (295-302)	1.3141
570	SDNLGFRC	SUMF2_HUMAN	Pos (283-290)	1.21
571	YINGKFKK	UCR11_HUMAN	Pos (47-54)	1.1939
572	EDNAAFYI	SO1C1_HUMAN	Pos (218-225)	1.3131
573	YPNERFEL	UBXD7_HUMAN	Pos (447-454)	1.2407
574	IYNISFEK	ZPBP1_HUMAN	Pos (185-192)	1.2866
575	YYNLPYTF	KV1C_HUMAN	Pos (91-98)	1.3132
576	IINATFFE	EMR3_HUMAN	Pos (248-255)	1.3126
577	TCNEGYDL	CSMD2_HUMAN	Pos (368-375)	1.1919
578	LYNLGFGN	ITIH3_HUMAN	Pos (418-425)	1.3122
579	LVNRGFVP	SURF1_HUMAN	Pos (176-183)	1.2823
580	IDNGNYDI	PRIC1_HUMAN	Pos (638-645)	1.3112
581	IENIAFGS	NAMPT_HUMAN	Pos (375-382)	1.3113
582	YDNAEYAE	CD109_HUMAN	Pos (655-662)	1.3109
583	LKNLAFME	CP2CJ_HUMAN	Pos (234-241)	1.2802
584	YANGCFEF	BIRC6_HUMAN	Pos (4593-4600)	1.1691
585	RWNDAFCD	FCER2_HUMAN	Pos (267-274)	1.1384
586	YMNRTFTD	KIT_HUMAN	Pos (350-357)	1.2802
587	KMNVAFEE	HM20B_HUMAN	Pos (208-215)	1.2333
588	ACNAAFPD	BIG1_HUMAN	Pos (1331-1338)	1.1886
		BIG2_HUMAN	Pos (1278-1285)	
589	YVNYQFRV	NFASC_HUMAN	Pos (695-702)	1.2884

590	YRNGKYDL	ENOB_HUMAN	Pos (251-258)	1.2057
591	VDNDGFIN	EFHB_HUMAN	Pos (609-616)	1.3084
592	RVNLGYLQ	KCNT1_HUMAN	Pos (1155-1162)	1.2324
593	YVNSSFYK	ABCG2_HUMAN	Pos (336-343)	1.2814
594	IINKVFEE	RASL2_HUMAN	Pos (371-378)	1.2784
595	GDNKAFNE	DGLA_HUMAN	Pos (632-639)	1.2783
596	INNAGFPA	FLRT2_HUMAN	Pos (75-82)	1.3077
597	LENDAFSD	OSBL5_HUMAN	Pos (295-302)	1.3073
598	INNRAYCH	PAP2_HUMAN	Pos (1745-1752)	1.1817
599	SVNHGFEA	SMAD5_HUMAN	Pos (395-402)	1.3068
600	KSNLAYDI	HSP74_HUMAN	Pos (84-91)	1.2305
601	HYNGTFED	FKB10_HUMAN	Pos (68-75)	1.3061
602	RQNVAYEY	IQGA1_HUMAN	Pos (36-43)	1.2301
603	YDNYIFWS	LRP1B_HUMAN	Pos (2414-2421)	1.3059
604	YANSGYSA	BET1_HUMAN	Pos (20-27)	1.3059
605	IANPAFVV	DRD2_HUMAN	Pos (184-191)	1.3056
606	VINSGFIC	OR4CF_HUMAN	Pos (199-206)	1.2221
607	IDNADFEI	IRK2_HUMAN	Pos (287-294)	1.3052
608	KDNPAFSM	GP178_HUMAN	Pos (569-576)	1.2291
609	LVNDAYKT	HSC20_HUMAN	Pos (123-130)	1.2829
610	LINVGFLE	B4GT7_HUMAN	Pos (144-151)	1.3032
611	KVNRAFNS	ACO11_HUMAN	Pos (114-121)	1.1978
612	VENEAFPK	CE110_HUMAN	Pos (182-189)	1.2762
613	IRNDAFLQ	GPR98_HUMAN	Pos (598-605)	1.2723
614	EKNEAFLD	GIPR_HUMAN	Pos (122-129)	1.2719
615	LLNYAYNG	USP9Y_HUMAN	Pos (1384-1391)	1.3023
616	CVNSAFRC	PLXA2_HUMAN	Pos (680-687)	1.0611
617	FDNGLYDE	CN102_HUMAN	Pos (667-674)	1.3011
618	LINQGYRA	CT059_HUMAN	Pos (300-307)	1.2806
619	YLNLMFND	LRC34_HUMAN	Pos (142-149)	1.3011
620	WLNVAYLD	OCTC_HUMAN	Pos (85-92)	1.301
621	YDNTSFGL	THSD1_HUMAN	Pos (812-819)	1.3009
622	HVNLGYTK	NPS3A_HUMAN	Pos (172-179)	1.2742
		NPS3B_HUMAN	Pos (172-179)	
623	DVNYAFLH	DCE2_HUMAN	Pos (86-93)	1.3002
624	INNRAFCN	PAPPA_HUMAN	Pos (1571-1578)	1.1749
625	VVNTLYED	DUS18_HUMAN	Pos (56-63)	1.2995
626	LENVGFEV	CO6_HUMAN	Pos (236-243)	1.2993
627	IRNLTFQD	GI24_HUMAN	Pos (89-96)	1.2689
628	YKNATYGY	CALU_HUMAN	Pos (129-136)	1.2686
629	VFNYGYLT	PIGB_HUMAN	Pos (98-105)	1.2982

630	ITNPGFGG	TROP_HUMAN	Pos (1050-1057)	1.2979
631	ILNGTYRD	QKI_HUMAN	Pos (199-206)	1.2769
632	FTNQAYAI	ADRB2_HUMAN	Pos (194-201)	1.2971
633	SDNFGYNL	AT8A2_HUMAN	Pos (316-323)	1.2967
634	YVNYVFQI	HERC3_HUMAN	Pos (893-900)	1.2967
635	MDNAGYSA	PCDGC_HUMAN	Pos (322-329)	1.2961
636	VVNYKYEQ	DOC10_HUMAN	Pos (122-129)	1.2231
637	YDNHTYGK	MYT1L_HUMAN	Pos (651-658)	1.269
638	YVNFRFPA	PABP5_HUMAN	Pos (63-70)	1.223
639	YVNQSFAP	ATG12_HUMAN	Pos (103-110)	1.2953
640	RDNQGFCI	S27A5_HUMAN	Pos (478-485)	1.1232
641	NVNGAFAE	LYL1_HUMAN	Pos (152-159)	1.2948
		TAL1_HUMAN	Pos (202-209)	
642	TVNKGYWF	OX2G_HUMAN	Pos (228-235)	1.2651
643	FSNCSFED	ADAM2_HUMAN	Pos (351-358)	1.1946
644	KPNNAYEF	SPAG1_HUMAN	Pos (802-809)	1.2187
645	SLNAAFEI	DEPD5_HUMAN	Pos (1348-1355)	1.2942
646	RLNDGFKD	S12A3_HUMAN	Pos (919-926)	1.1972
647	IINSSFEN	CENPF_HUMAN	Pos (2509-2516)	1.2938
648	SENLAFID	KCNK4_HUMAN	Pos (341-348)	1.2935
649	RENIAFEL	IFNK_HUMAN	Pos (61-68)	1.2175
650	IGNSAFAR	VPS39_HUMAN	Pos (859-866)	1.2666
651	IHNCAFNG	FSHR_HUMAN	Pos (185-192)	1.1931
652	LVNPAFPG	ZP2_HUMAN	Pos (44-51)	1.293
653	ICNRGFSS	PATZ1_HUMAN	Pos (499-506)	1.1428
654	IINAIYEP	NP1L1_HUMAN	Pos (120-127)	1.2926
		NP1L2_HUMAN	Pos (154-161)	
655	QENPAYEE	ACOX2_HUMAN	Pos (661-668)	1.2924
656	ICNTGYSL	FHR5_HUMAN	Pos (113-120)	1.1709
657	VLNLAFAQ	ANC5_HUMAN	Pos (624-631)	1.2912
658	IINECYNY	K1333_HUMAN	Pos (513-520)	1.15
659	RQNIAYEY	IQGA2_HUMAN	Pos (32-39)	1.2151
660	LVNNAYRG	CI096_HUMAN	Pos (562-569)	1.2703
661	SDNEGYFL	PRIC2_HUMAN	Pos (784-791)	1.2903
662	QINRAYQF	K0103_HUMAN	Pos (263-270)	1.2608
663	HKNQGYYD	ODO1_HUMAN	Pos (960-967)	1.2591
664	FGNGAYTM	RN139_HUMAN	Pos (480-487)	1.2884
665	KYNEAYIS	CSR2B_HUMAN	Pos (695-702)	1.2124
666	LSNYAFLY	GP155_HUMAN	Pos (313-320)	1.2878
667	LVNNAYAG	DHRS1_HUMAN	Pos (89-96)	1.2875
668	LRNSAFES	U520_HUMAN	Pos (1312-1319)	1.2569

669	LLNYAYNS	USP9X_HUMAN	Pos (1375-1382)	1.2868
670	LDNPGYSV	KCNS2_HUMAN	Pos (177-184)	1.2864
671	YPNPTFEL	PLXA2_HUMAN	Pos (1155-1162)	1.2863
672	GVNCTFED	PLXA4_HUMAN	Pos (590-597)	1.186
673	VINESFEG	ROS_HUMAN	Pos (2238-2245)	1.2859
674	RINDAFRL	ACE2_HUMAN	Pos (710-717)	1.1888
675	SVNHGFET	SMAD1_HUMAN	Pos (395-402)	1.2845
676	RENEGYYF	CD8A_HUMAN	Pos (107-114)	1.2078
677	IDNPDYYD	MFAP2_HUMAN	Pos (42-49)	1.2834
678	ILNTGFTI	FREM2_HUMAN	Pos (832-839)	1.2832
679	YINHLYSD	PLSI_HUMAN	Pos (413-420)	1.2832
680	VLNPAFRE	GNAS1_HUMAN	Pos (80-87)	1.2624
681	SDNAGFCI	SCRB2_HUMAN	Pos (305-312)	1.1868
682	LINSAFLK	INT7_HUMAN	Pos (60-67)	1.2557
683	LVNRGYQP	GCP6_HUMAN	Pos (1794-1801)	1.2527
684	PTNLAYDG	ENK19_HUMAN	Pos (290-297)	1.2822
685	IWNGTFEE	AT10B_HUMAN	Pos (337-344)	1.2816
686	FCNDGYRL	CSMD3_HUMAN	Pos (2620-2627)	1.1402
687	DINVGFDT	IRK5_HUMAN	Pos (256-263)	1.2812
688	LVNASFKD	PA24D_HUMAN	Pos (732-739)	1.2602
689	LQNDAYND	CF113_HUMAN	Pos (373-380)	1.2808
690	LENSGFDR	UBP37_HUMAN	Pos (695-702)	1.2536
691	LGNPGFDE	GRK4_HUMAN	Pos (279-286)	1.2801
692	QYNGTYDD	LRFN5_HUMAN	Pos (450-457)	1.2791
693	VFNKGYGF	VDAC3_HUMAN	Pos (17-24)	1.2495
694	ILNLTFSD	CAD13_HUMAN	Pos (50-57)	1.2785
695	EVNEAFAP	THIM_HUMAN	Pos (318-325)	1.2784
696	IRNQGYST	WT1_HUMAN	Pos (144-151)	1.2476
697	FSNRAFAL	TTL12_HUMAN	Pos (495-502)	1.2484
698	ITNKMFED	SP100_HUMAN	Pos (84-91)	1.2483
699	FENYSYDL	GPR1_HUMAN	Pos (12-19)	1.2777
700	KLNPGFDL	PLAL1_HUMAN	Pos (292-299)	1.2016
701	YDNQVYWY	NCKX5_HUMAN	Pos (191-198)	1.2772
702	YDNFLFLC	PAEP_HUMAN	Pos (117-124)	1.1938
703	VCNKAYKR	ZN319_HUMAN	Pos (432-439)	1.0799
704	YINGRYYS	CD166_HUMAN	Pos (464-471)	1.2044
705	LQNSGFEY	RNF14_HUMAN	Pos (58-65)	1.2769
706	YDNLNYTL	3BHS1_HUMAN	Pos (264-271)	1.2767
707	TNNEAYFD	AP3M1_HUMAN	Pos (175-182)	1.2762
		AP3M2_HUMAN	Pos (175-182)	
708	GFNFGFDL	GMCL1_HUMAN	Pos (413-420)	1.2759

		GMCLL_HUMAN	Pos (413-420)	
709	DINVGFDS	IRK2_HUMAN	Pos (249-256)	1.276
710	YDNVVYRY	SO3A1_HUMAN	Pos (617-624)	1.2553
711	HDNTAFLE	TITIN_HUMAN	Pos (5653-5660)	1.2756
712	RINLGFSN	DMBT1_HUMAN	Pos (1803-1810)	1.1996
713	FDNNCYRD	SC6A2_HUMAN	Pos (335-342)	1.1131
714	IINAEYEP	NP1L3_HUMAN	Pos (150-157)	1.275
715	RDNKGFFN	CALL6_HUMAN	Pos (83-90)	1.1695
716	YVNYNFSN	BPIL2_HUMAN	Pos (75-82)	1.2745
717	ARNLAYDT	PLOD1_HUMAN	Pos (228-235)	1.2439
718	LRNYGYRH	HPLN4_HUMAN	Pos (249-256)	1.2231
719	PRNLGFET	PI4KB_HUMAN	Pos (698-705)	1.2431
720	FENSAYVV	GSH1_HUMAN	Pos (435-442)	1.2733
721	GVNRAYPY	CEBPZ_HUMAN	Pos (489-496)	1.2437
722	FINKCYEL	MTM1_HUMAN	Pos (187-194)	1.1025
723	YDNDSYTL	ASB12_HUMAN	Pos (35-42)	1.273
724	LVNNGYVR	LCAT_HUMAN	Pos (152-159)	1.2464
725	SDNGGFKR	TSSC4_HUMAN	Pos (117-124)	1.225
726	QFNQGFYD	DDX56_HUMAN	Pos (303-310)	1.2718
727	GYNAAYNR	MGP_HUMAN	Pos (87-94)	1.2451
728	MVNYSYDE	NR5A2_HUMAN	Pos (73-80)	1.2709
729	EINEAFAA	THIC_HUMAN	Pos (319-326)	1.2709
730	VLNRAFKK	SYNJ2_HUMAN	Pos (312-319)	1.1938
731	CNNLGYEI	LIPL_HUMAN	Pos (310-317)	1.1336
732	PPNLGFVD	MLL2_HUMAN	Pos (2026-2033)	1.2695
733	IINKLFFD	TRAIP_HUMAN	Pos (58-65)	1.2398
734	KENVAFEQ	LACTB_HUMAN	Pos (243-250)	1.1928
735	GLNQAFEE	SYNE1_HUMAN	Pos (6357-6364)	1.2678
736	RVNLSFDF	PXDC2_HUMAN	Pos (158-165)	1.1919
737	LYNTGYYL	CD109_HUMAN	Pos (616-623)	1.2675
738	YYNEKYGG	CBPN_HUMAN	Pos (171-178)	1.1947
739	VCNAGFHV	FBN1_HUMAN	Pos (556-563)	1.1465
740	YVNWRFMR	SMUF1_HUMAN	Pos (591-598)	1.1676
741	GVNAAFQA	MTMR5_HUMAN	Pos (287-294)	1.2654
742	RTNSGFEE	ANR44_HUMAN	Pos (554-561)	1.1892
743	YDNLNYIL	3BHS2_HUMAN	Pos (263-270)	1.2643
744	VANLAFAK	MYO15_HUMAN	Pos (1294-1301)	1.2377
745	FWNKAFLT	BAP29_HUMAN	Pos (47-54)	1.2347
746	LINSGFPE	DHX9_HUMAN	Pos (964-971)	1.2641
747	LTNYGFYG	PA2G5_HUMAN	Pos (38-45)	1.2641
748	ISNVKFED	NVL_HUMAN	Pos (259-266)	1.1914

749	ISNVRYED	FSTL5_HUMAN	Pos (400-407)	1.1914
750	RSNAGFDG	LRRC6_HUMAN	Pos (201-208)	1.1879
751	SENGAYLD	CTGE5_HUMAN	Pos (299-306)	1.2634
752	TYNLAFCG	OR5M8_HUMAN	Pos (162-169)	1.1673
753	VENEGFTL	ANKF1_HUMAN	Pos (232-239)	1.2631
754	YNNITFEE	CSN4_HUMAN	Pos (316-323)	1.2626
755	GVNCSFED	PLXA1_HUMAN	Pos (595-602)	1.1621
756	WDNPAYSG	PK2L1_HUMAN	Pos (20-27)	1.2617
757	YANLCFEA	LCTL_HUMAN	Pos (181-188)	1.1203
758	ACNRGYYD	ANR11_HUMAN	Pos (209-216)	1.1114
759	ISNMAFLV	O13C3_HUMAN	Pos (201-208)	1.2612
760	YDNGRFIG	SV2C_HUMAN	Pos (497-504)	1.1888
761	IDNYCFFQ	KLH11_HUMAN	Pos (543-550)	1.1198
762	YDNHPFWF	MBOA5_HUMAN	Pos (261-268)	1.2605
763	RVNGGFSG	MA1A2_HUMAN	Pos (566-573)	1.1843
764	RVNGGYSG	MA1A1_HUMAN	Pos (580-587)	1.1843
765	IQNNAYAV	GBRA2_HUMAN	Pos (362-369)	1.2596
766	LCNFAFYG	CTL5_HUMAN	Pos (437-444)	1.1392
767	VGNHAYEN	MCLN3_HUMAN	Pos (141-148)	1.2595
768	FLNMAYQG	CD1E_HUMAN	Pos (127-134)	1.2593
769	HGNFAFEG	SNX17_HUMAN	Pos (457-464)	1.2593
770	FDNETFRI	BMP7_HUMAN	Pos (185-192)	1.2384
771	ITNNGFNP	PLCD1_HUMAN	Pos (675-682)	1.2589
772	VLNEGFLV	B3GN1_HUMAN	Pos (365-372)	1.2583
773	WKNAAFLY	UTX_HUMAN	Pos (127-134)	1.2279
		UTY_HUMAN	Pos (124-131)	
774	IDNKMYDP	ENPP1_HUMAN	Pos (275-282)	1.2286
775	FENVGFTK	MSS4_HUMAN	Pos (77-84)	1.2312
776	YENGSFYF	NEUA_HUMAN	Pos (212-219)	1.2575
777	QLNKAYEA	TANK_HUMAN	Pos (8-15)	1.2279
778	WLNNAFQD	FBXW5_HUMAN	Pos (219-226)	1.2571
779	QVNEGFQP	CAH4_HUMAN	Pos (175-182)	1.2564
780	FDNNSFEQ	MYO1D_HUMAN	Pos (390-397)	1.2563
781	YINANYMP	PTPRJ_HUMAN	Pos (1094-1101)	1.2562
782	VYNGTYGD	MUC2_HUMAN	Pos (4349-4356)	1.2559
783	LENSGFEP	PGS1_HUMAN	Pos (195-202)	1.2554
784	FGNAGFSV	ASPH1_HUMAN	Pos (263-270)	1.2549
785	IINYSYIN	UBP24_HUMAN	Pos (96-103)	1.2548
786	TKNSAFEE	K0232_HUMAN	Pos (699-706)	1.2243
787	LINNAFQL	ARTS1_HUMAN	Pos (627-634)	1.2547
788	FPNLGYIQ	FA38A_HUMAN	Pos (437-444)	1.2545

789	FENPAFSG	CRBB3_HUMAN	Pos (119-126)	1.2544
790	LENYAYPG	ACON_HUMAN	Pos (178-185)	1.2538
791	LLNYGYVG	NRX1A_HUMAN	Pos (635-642)	1.2538
792	VVNLKFLD	LRRN3_HUMAN	Pos (259-266)	1.1811
793	VKNLGFGG	CPLX4_HUMAN	Pos (13-20)	1.2233
794	VCNLAYSG	PSD10_HUMAN	Pos (10-17)	1.1329
795	TVNRAYGL	O52L1_HUMAN	Pos (196-203)	1.2229
		O52L2_HUMAN	Pos (196-203)	
796	KDNQGYTP	ANR44_HUMAN	Pos (768-775)	1.1762
797	VKNKAYFK	RL5_HUMAN	Pos (6-13)	1.1655
798	LVNARFDY	TEN1_HUMAN	Pos (2034-2041)	1.1793
799	VRNFGYPL	TAB1_HUMAN	Pos (362-369)	1.2209
800	SLNDAYDF	DUS7_HUMAN	Pos (302-309)	1.2513
801	FDNLDFDL	ATF6A_HUMAN	Pos (57-64)	1.2509
802	TRNVAFER	AN13D_HUMAN	Pos (150-157)	1.1933
803	VTNCAFSF	CN104_HUMAN	Pos (312-319)	1.1498
804	LINDGFQR	CASZ1_HUMAN	Pos (459-466)	1.2231
805	IINHYFCD	OR8D1_HUMAN	Pos (173-180)	1.1536
806	YVNVFYKF	ZN430_HUMAN	Pos (182-189)	1.2287
807	YFNYKYGW	CCG5_HUMAN	Pos (170-177)	1.1768
808	ALNKAYHD	MACF1_HUMAN	Pos (1513-1520)	1.2193
809	IENLSYDA	SACS_HUMAN	Pos (2735-2742)	1.2486
810	FGNKAYRT	PTPA_HUMAN	Pos (135-142)	1.1984
811	LLNEAFVV	CHLE_HUMAN	Pos (301-308)	1.2482
812	TNNLGYMD	CAND1_HUMAN	Pos (767-774)	1.2482
813	QDNVAFVL	NCLN_HUMAN	Pos (302-309)	1.2479
814	IINQSFKF	K1377_HUMAN	Pos (594-601)	1.2268
815	FVNAVYDT	HECD2_HUMAN	Pos (189-196)	1.2471
816	FSNGAFLG	SIG10_HUMAN	Pos (548-555)	1.247
817	NVNYGYQQ	ALK_HUMAN	Pos (1581-1588)	1.247
818	YFNEEFDL	AN13A_HUMAN	Pos (335-342)	1.2469
819	EVNFGYIE	F10C1_HUMAN	Pos (190-197)	1.2467
820	MENQAFDP	TRAM2_HUMAN	Pos (274-281)	1.2466
821	EYNYGFVV	F125A_HUMAN	Pos (253-260)	1.2464
822	FLNMAYQQ	STEA2_HUMAN	Pos (333-340)	1.2464
823	NCNEGFEP	FBN2_HUMAN	Pos (2237-2244)	1.1259
824	AVNVGFCF	NSD3_HUMAN	Pos (859-866)	1.1502
825	VFNWAFKK	SIG11_HUMAN	Pos (166-173)	1.1987
826	YQNFVFEE	LYST_HUMAN	Pos (696-703)	1.2458
827	KENEGFGF	MAGI2_HUMAN	Pos (926-933)	1.1698
828	EVNEAFSL	THIL_HUMAN	Pos (351-358)	1.245

829	IVNLRYFE	CAC1E_HUMAN	Pos (1148-1155)	1.1726
830	SWNFGFEV	JPH2_HUMAN	Pos (41-48)	1.2452
831	TKNYGYVY	CF206_HUMAN	Pos (255-262)	1.2143
832	LHNQAFEI	RBCC1_HUMAN	Pos (1161-1168)	1.2442
833	YYNYFFVS	ASM3A_HUMAN	Pos (406-413)	1.2442
834	LHNVGFAD	OBSCN_HUMAN	Pos (2046-2053)	1.244
835	LRNAAYFL	CENPU_HUMAN	Pos (351-358)	1.2132
836	IYNFTYKV	ABCAC_HUMAN	Pos (646-653)	1.2229
837	KINETFVD	PPIL4_HUMAN	Pos (138-145)	1.1677
838	IGNSAFLL	MA2A1_HUMAN	Pos (599-606)	1.2434
839	EINEAFAS	THIK_HUMAN	Pos (343-350)	1.2434
840	YFNESFSF	SYT2_HUMAN	Pos (336-343)	1.2433
841	FVNGVFEV	IF16_HUMAN	Pos (686-693)	1.2431
842	YRNSSYEN	ZBT38_HUMAN	Pos (581-588)	1.2129
843	KKNASFED	OR2J1_HUMAN	Pos (4-11)	1.1369
844	FANAGFTS	EPHB2_HUMAN	Pos (934-941)	1.2428
845	FSNGGFFV	CSMD1_HUMAN	Pos (2241-2248)	1.2424
846	RTNCGYES	CF058_HUMAN	Pos (109-116)	1.0663
847	FKNNAYLA	SYDC_HUMAN	Pos (240-247)	1.2114
848	MANEGFDP	S27A3_HUMAN	Pos (641-648)	1.2413
849	LINKLYDD	UN45B_HUMAN	Pos (358-365)	1.2118
850	RVNSGFGR	STCH_HUMAN	Pos (340-347)	1.1388
851	DPNFGFEE	ATP4B_HUMAN	Pos (167-174)	1.2413
852	YHNTTFDW	S4A4_HUMAN	Pos (659-666)	1.241
853	KVNATFER	GCYA2_HUMAN	Pos (370-377)	1.1384
854	ITNLKYEE	MUC16_HUMAN	Pos (12554-12561)	1.1683
855	ITNLRYEE	MUC16_HUMAN	Pos (15363-15370)	1.1683
856	LVNASYTD	TRPV1_HUMAN	Pos (190-197)	1.2402
857	FVNPKFEH	KPCD_HUMAN	Pos (665-672)	1.1675
858	QDNTAFVL	RLF_HUMAN	Pos (291-298)	1.2398
859	VCNPGFLF	ZAN_HUMAN	Pos (1075-1082)	1.1194
860	QDNPAYLL	SNPC4_HUMAN	Pos (1356-1363)	1.2389
861	SYNAGYSY	ELOV4_HUMAN	Pos (101-108)	1.239
862	YINANFIK	PTN12_HUMAN	Pos (88-95)	1.2124
		PTN22_HUMAN	Pos (84-91)	
863	YINANYIK	PTN6_HUMAN	Pos (301-308)	1.2124
864	YINANYIR	PTN5_HUMAN	Pos (329-336)	1.2124
		PTN7_HUMAN	Pos (149-156)	
		PTPRR_HUMAN	Pos (446-453)	
865	FCNPGFPI	TM9S2_HUMAN	Pos (173-180)	1.1183
866	AVNYEFED	DAPK1_HUMAN	Pos (231-238)	1.2387

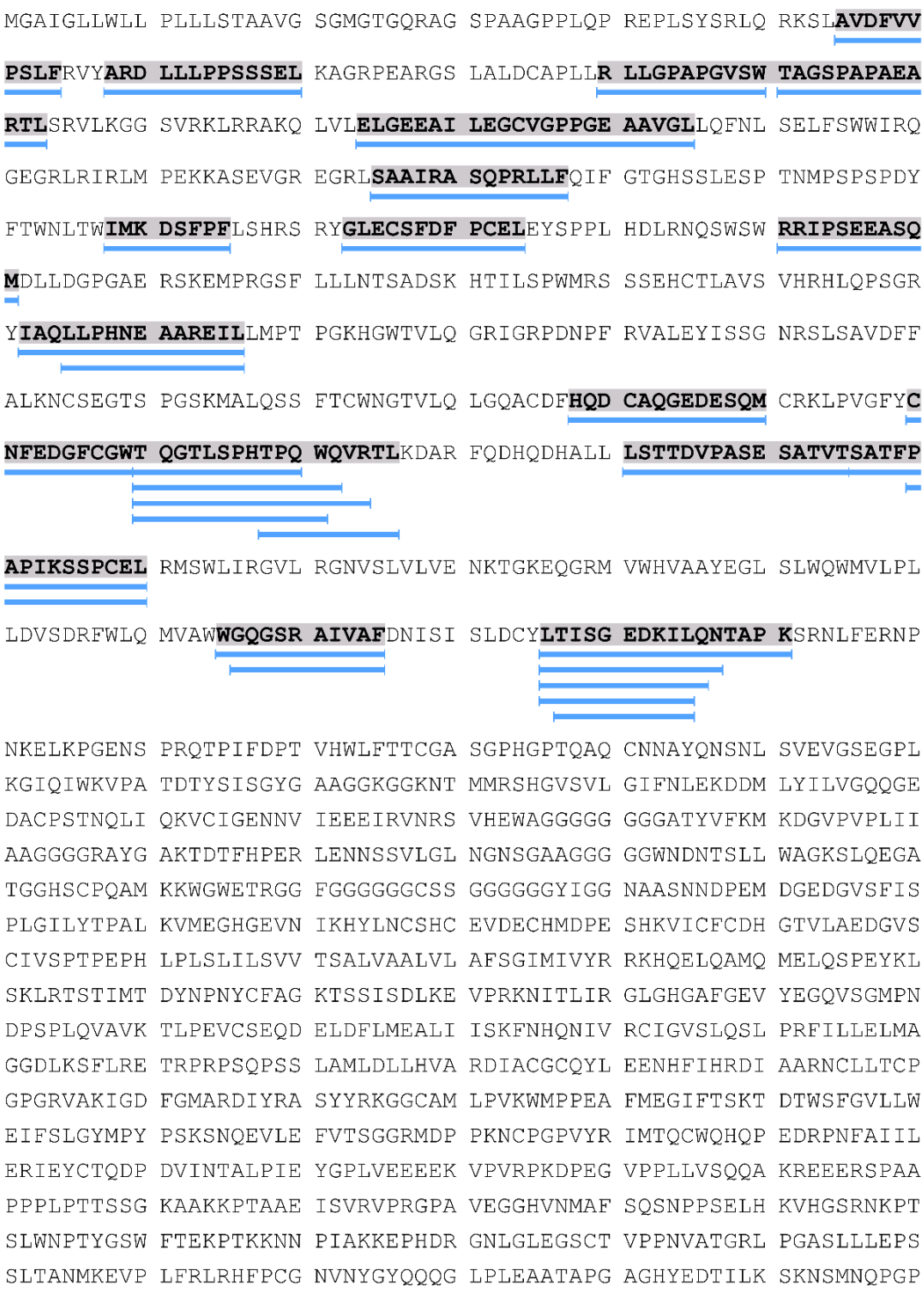
867	VFNRAFYS	MBTP2_HUMAN	Pos (56-63)	1.2092
868	NPNAAYDK	DCTN2_HUMAN	Pos (48-55)	1.212
869	RLNAAYSF	MKS3_HUMAN	Pos (363-370)	1.1625
870	TLNQAYQD	ABCA2_HUMAN	Pos (593-600)	1.2382
871	KVNTGFLM	CO038_HUMAN	Pos (90-97)	1.1623
872	LRNGSYED	ABCAC_HUMAN	Pos (410-417)	1.2076
873	CDNSAFII	ZN248_HUMAN	Pos (307-314)	1.1012
874	ISNWAFIP	GP115_HUMAN	Pos (256-263)	1.2378
875	FENQGFFQ	PASK_HUMAN	Pos (1071-1078)	1.2376
876	RINNAFFL	TMM27_HUMAN	Pos (111-118)	1.1615
877	YDNGYYLP	E2AK3_HUMAN	Pos (475-482)	1.2369
878	RRNVGFES	NUP50_HUMAN	Pos (45-52)	1.1307
879	WVNWGYAT	NUAK2_HUMAN	Pos (302-309)	1.2368
880	IKNSGYIS	E2AK3_HUMAN	Pos (579-586)	1.2063
881	VLNGAFLA	A4GAT_HUMAN	Pos (215-222)	1.2364
882	LENQAFCF	KIF2C_HUMAN	Pos (304-311)	1.1402
883	TKNIAFER	AN13A_HUMAN	Pos (238-245)	1.1783
884	WVNWGYKS	NUAK1_HUMAN	Pos (305-312)	1.2144
885	YVNALFSK	HXA7_HUMAN	Pos (6-13)	1.2081
886	YFNIPYDW	ERG25_HUMAN	Pos (122-129)	1.2348
887	LLNAGYKK	RPC1_HUMAN	Pos (716-723)	1.187
888	SYNGGYGF	KR191_HUMAN	Pos (79-86)	1.2341
889	PVNRAYFA	PPM1F_HUMAN	Pos (188-195)	1.2043
890	IDNRDFPD	PIGL_HUMAN	Pos (109-116)	1.2041
891	LDNDGYIT	NCS1_HUMAN	Pos (109-116)	1.2335
892	LENKGYYF	PCSK4_HUMAN	Pos (556-563)	1.2041
893	RINGTYDF	LPH_HUMAN	Pos (1670-1677)	1.1575
894	LYNRGFSL	ERBB3_HUMAN	Pos (423-430)	1.2036
895	SVNFGFSK	MRE11_HUMAN	Pos (98-105)	1.2065
896	VSNKAYVF	SCRB2_HUMAN	Pos (111-118)	1.2034
897	LKNAGFTE	LMTK2_HUMAN	Pos (749-756)	1.2022
898	LCNFAFKT	ZEP2_HUMAN	Pos (1831-1838)	1.0909
899	IRNLSFDS	RBM28_HUMAN	Pos (338-345)	1.2013
900	YNNYIYDL	KIF23_HUMAN	Pos (226-233)	1.2315
901	FINPIFEF	NR1H2_HUMAN	Pos (349-356)	1.2315
		NR1H3_HUMAN	Pos (335-342)	
902	AINAAFSA	MS3L1_HUMAN	Pos (218-225)	1.2313
903	LHNLAFAF	SPTN5_HUMAN	Pos (235-242)	1.2312
904	FDNVHFEY	ABCB7_HUMAN	Pos (474-481)	1.231
905	YLNETFSE	PYR1_HUMAN	Pos (1558-1565)	1.2309
906	PTNYGFCF	GRB14_HUMAN	Pos (301-308)	1.1348

907	IVNHFFCD	O5AS1_HUMAN	Pos (173-180)	1.1338
908	IVNHFYCD	OR8U1_HUMAN	Pos (173-180)	1.1338
909	VVNELFRD	BCLX_HUMAN	Pos (126-133)	1.2089
910	YSNAYFTD	SEM5A_HUMAN	Pos (1060-1067)	1.2296
911	FTNIAYFT	AAA1_HUMAN	Pos (282-289)	1.2294
912	VPNLGFGM	LIPP_HUMAN	Pos (227-234)	1.2288
913	IPNNAYLH	IFIT2_HUMAN	Pos (277-284)	1.2285
914	DENEGYEH	ZFYV9_HUMAN	Pos (370-377)	1.2284
915	GLNGGYQD	MTG8R_HUMAN	Pos (325-332)	1.2281
916	VCNKAFRT	ZN295_HUMAN	Pos (941-948)	1.0576
917	VCNHAFHF	RBX1_HUMAN	Pos (74-81)	1.1075
918	KDNIGFHH	MEPE_HUMAN	Pos (42-49)	1.1519
919	FWNNAFTL	NOC4L_HUMAN	Pos (206-213)	1.2276
920	NVNLAYLG	RNPC1_HUMAN	Pos (82-89)	1.2274
921	QVNSSFED	ATR_HUMAN	Pos (577-584)	1.2272
922	LVNMGFPQ	CJ035_HUMAN	Pos (71-78)	1.2269
923	YDNGNFFP	HDAC5_HUMAN	Pos (895-902)	1.2267
924	EENEGFEH	CJ086_HUMAN	Pos (336-343)	1.2267
925	KINCNFED	ENTK_HUMAN	Pos (344-351)	1.0508
926	ITNERFGD	SRBS1_HUMAN	Pos (772-779)	1.1535
927	ACNVGYYD	ANR12_HUMAN	Pos (225-232)	1.1053
928	IANTAYLQ	GRK7_HUMAN	Pos (11-18)	1.2257
929	YVNWRFLR	SMUF2_HUMAN	Pos (585-592)	1.1263
930	LVNYVYRD	TRRAP_HUMAN	Pos (2420-2427)	1.2047
931	FVNHCYVD	ITPR2_HUMAN	Pos (1426-1433)	1.084
		ITPR3_HUMAN	Pos (1416-1423)	
932	IDNPMFDT	MPP5_HUMAN	Pos (99-106)	1.2249
933	FNNRAYSK	SPF30_HUMAN	Pos (181-188)	1.1689
934	HVNPAFFP	CPSF6_HUMAN	Pos (356-363)	1.2248
935	YFNTRYPD	S13A1_HUMAN	Pos (268-275)	1.1523
936	YSNFRFDT	SIDT1_HUMAN	Pos (568-575)	1.1521
937	FPNFTFRD	ENPP7_HUMAN	Pos (265-272)	1.2036
938	ITNTTFDF	AF9_HUMAN	Pos (538-545)	1.2239
939	GINRAYGK	AUMH_HUMAN	Pos (93-100)	1.1673
940	EKNQGFDF	MK06_HUMAN	Pos (520-527)	1.1928
941	FINSPFRD	TRPV4_HUMAN	Pos (226-233)	1.2025
942	YTNYSFTL	USH2A_HUMAN	Pos (3651-3658)	1.2229
943	HTNQAFEI	LIPL_HUMAN	Pos (357-364)	1.2228
944	HENRGFEI	RNF6_HUMAN	Pos (157-164)	1.1933
945	YKNYSYLH	CHD8_HUMAN	Pos (393-400)	1.1923
		CHD9_HUMAN	Pos (720-727)	

946	HVNYTFEI	EPHA6_HUMAN	Pos (407-414)	1.2226
947	FSNSGFKE	P3C2A_HUMAN	Pos (5-12)	1.2016
948	IDNYLYVV	BTBD5_HUMAN	Pos (524-531)	1.2221
949	YSNLVYDW	GUP1_HUMAN	Pos (293-300)	1.2222
950	LVNYPFID	ERF_HUMAN	Pos (113-120)	1.2215
951	MVNLQFED	MASP1_HUMAN	Pos (223-230)	1.2215
952	INNPAFIS	SGTB_HUMAN	Pos (219-226)	1.2215
953	YKNYRYRA	SMO_HUMAN	Pos (394-401)	1.0976
954	GVNPAYPE	ETS1_HUMAN	Pos (144-151)	1.2207
955	WPNLAYWF	AMPD3_HUMAN	Pos (450-457)	1.2208
956	LGNLAFAE	TTC16_HUMAN	Pos (378-385)	1.2204
957	DVNVTFED	PFD4_HUMAN	Pos (12-19)	1.2202
958	LFNLGFVC	TUSC4_HUMAN	Pos (82-89)	1.1368
959	PSNLGYEL	PK1L1_HUMAN	Pos (396-403)	1.2194
960	YKNYFFVF	AT133_HUMAN	Pos (1102-1109)	1.1891
961	QCNQGYEL	FBLN3_HUMAN	Pos (278-285)	1.0987
962	FRNLTFEG	GUC2C_HUMAN	Pos (355-362)	1.1887
963	VFNPAFYS	T2R44_HUMAN	Pos (70-77)	1.2187
964	QPNSAYDF	SMC2_HUMAN	Pos (950-957)	1.2186
965	KYNPGYVL	XRN1_HUMAN	Pos (910-917)	1.1424
966	SSNRAYDF	AAT1_HUMAN	Pos (32-39)	1.1887
967	PINFVFED	TITIN_HUMAN	Pos (15732-15739)	1.218
968	GGNFGFGD	ROA2_HUMAN	Pos (204-211)	1.2178
969	FWNSAYQA	VGLL2_HUMAN	Pos (145-152)	1.2178
970	ITNHGFLI	TRIM5_HUMAN	Pos (454-461)	1.2175
971	FMNQAFLQ	DZIP1_HUMAN	Pos (207-214)	1.2174
972	ILNRTFQD	TIP_HUMAN	Pos (149-156)	1.1877
973	LGNYAYCQ	FANCM_HUMAN	Pos (220-227)	1.1212
974	YANYTFNV	EPHA1_HUMAN	Pos (412-419)	1.2168
975	YGNVPFEY	GPTC3_HUMAN	Pos (232-239)	1.2166
976	EFNKAFEL	CALB1_HUMAN	Pos (189-196)	1.1869
977	VLNGGFRN	THTR_HUMAN	Pos (125-132)	1.1957
978	ILNHVFDD	ES8L1_HUMAN	Pos (260-267)	1.2161
979	FVNLTFTK	LAMP3_HUMAN	Pos (289-296)	1.1894
980	HINPAYIK	TTC6_HUMAN	Pos (275-282)	1.1893
981	WRNYGFGI	CHSS1_HUMAN	Pos (688-695)	1.1855
982	VCNAGYSG	TENX_HUMAN	Pos (451-458)	1.0954
983	PVNNAFSY	MCR_HUMAN	Pos (336-343)	1.2152
984	STNCAYDR	DNSL3_HUMAN	Pos (228-235)	1.0885
985	LKNEGFYS	LAT3_HUMAN	Pos (40-47)	1.1846
986	NVNLGFLL	YSK4_HUMAN	Pos (154-161)	1.2145

987	SDNVGFGE	WNT8B_HUMAN	Pos (133-140)	1.2146
988	FDNIKYDN	GLTL4_HUMAN	Pos (280-287)	1.1419
989	EVNKAYWR	RM39_HUMAN	Pos (137-144)	1.1582
990	FDNIRFDT	GALT8_HUMAN	Pos (307-314)	1.1414
991	VLNPAFYS	T2R50_HUMAN	Pos (70-77)	1.2138
992	YENETFLC	MOB3_HUMAN	Pos (181-188)	1.1303
993	YINWKYFL	DJBP_HUMAN	Pos (1124-1131)	1.1409
994	IINHFFCD	OR5A1_HUMAN	Pos (176-183)	1.1172
		OR6A2_HUMAN	Pos (178-185)	
		OR6J1_HUMAN	Pos (172-179)	
		OR6P1_HUMAN	Pos (173-180)	
995	IINHFYCD	OR8J1_HUMAN	Pos (173-180)	1.1172
		OR8J3_HUMAN	Pos (173-180)	
996	PDNTAFKQ	CC50A_HUMAN	Pos (28-35)	1.1923
997	RKNEGYMR	IF38_HUMAN	Pos (893-900)	1.08
998	VINLNYKD	SSRA_HUMAN	Pos (172-179)	1.1923
999	VVNELYVD	ERGI1_HUMAN	Pos (52-59)	1.2127
1000	VNNLAFLK	ABCAD_HUMAN	Pos (109-116)	1.186

S7 MS analysis of ALK ectodomain (chymotrypsin digestion)



S8 MS analysis of ALK ectodomain (Glu-C digestion)

```
   1  MGAIGLLWLL PLLLSTAAVG SGMGTGQRAG SPAAGPPLQP REPLSYSRLQ RKSLAVDFVV
  61  PSLFRVYARD LLLPPSSSEL KAGRPEARGS LALDCAPLLR LLGPAPGVSW TAGSPAPAEA
 121  RTLSRVLKGG SVRKLRRAKQ LVLELGEEAI LEGCVGPPGE AAVGLLQFNL SELFSWWIRQ

 181  GEGRLRIRLM PEKKASEVGR EGRLSAAIRA SQPRLLFQIF GTGHSSLESP TNMPSPSPDY

 241  FTWNLTWIMK DSFPFLSHRS RYGLECSFDF PCELEYSPPL HDLRNQSWSW RRIPSEEASQ

 301  MDLLDGPGAE RSKEMPRGSF LLLNTSADSK HTILSPWMRS SSEHCTLAVS VHRHLQPSGR
 361  YIAQLLPHNE AAREILLMPT PGKHGWTVLQ GRIGRPDNPF RVALEYISSG NRSLSAVDFF
 421  ALKNCSEGTS PGSKMALQSS FTCWNGTVLQ LGQACDFHQD CAQGEDESQM CRKLPVGFYC

 481  NFEDGFCGWT QGTLSPHTPQ WQVRTLKDAR FQDHQDHALL LSTTDVPASE SATVTSATFP

 541  APIKSSPCEL RMSWLIRGVL RGNVSLVLVE NKTGKEQGRM VWHVAAYEGL SLWQWMVLPL
 601  LDVSDRFWLQ MVAWWGQGSR AIVAFDNISI SLDCYLTISG EDKILQNTAP KSRNLFERNP

 661  NKELKPGENS PRQTPIFDPT VHWLFTTCGA SGPHGPTQAQ CNNAYQNSNL SVEVGSEGPL
 721  KGIQIWKVPA TDTYSISGYG AAGGKGGKNT MMRSHGVSVL GIFNLEKDDM LYILVGQQGE
 781  DACPSTNQLI QKVCIGENNV IEEEIRVNRS VHEWAGGGGG GGGATYVFKM KDGVPVPLII
 841  AAGGGGRAYG AKTDTFHPER LENNSSVLGL NGNSGAAGGG GGWNDNTSLL WAGKSLQEGA
 901  TGGHSCPQAM KKWGWETRGG FGGGGGGCSS GGGGGGYIGG NAASNNDPEM DGEDGVSFIS
 961  PLGILYTPAL KVMEGHGEVN IKHYLNCSHC EVDECHMDPE SHKVICFCDH GTVLAEDGVS
1021  CIVSPTPEPH LPLSLILSVV TSALVAALVL AFSGIMIVYR RKHQELQAMQ MELQSPEYKL
1081  SKLRTSTIMT DYNPNYCFAG KTSSISDLKE VPRKNITLIR GLGHGAFGEV YEGQVSGMPN
1141  DPSPLQVAVK TLPEVCSEQD ELDFLMEALI ISKFNHQNIV RCIGVSLQSL PRFILLELMA
1201  GGDLKSFLRE TRPRPSQPSS LAMLDLLHVA RDIACGCQYL EENHFIHRDI AARNCLLTCP
1261  GPGRVAKIGD FGMARDIYRA SYYRKGGCAM LPVKWMPPEA FMEGIFTSKT DTWSFGVLLW
1321  EIFSLGYMPY PSKSNQEVLE FVTSGGRMDP PKNCPGPVYR IMTQCWQHQP EDRPNFAIIL
1381  ERIEYCTQDP DVINTALPIE YGPLVEEEEK VPVRPKDPEG VPPLLVSQQA KREEERSPAA
1441  PPPLPTTSSG KAAKKPTAAE ISVRVPRGPA VEGGHVNMAF SQSNPPSELH KVHGSRNKPT
1501  SLWNPTYGSW FTEKPTKKNN PIAKKEPHDR GNLGLEGSCT VPPNVATGRL PGASLLLEPS
1561  SLTANMKEVP LFRLRHFPCG NVNYGYQQQG LPLEAATAPG AGHYEDTILK SKNSMNQPGP
```

S9 MS analysis of ALK ectodomain (Trypsin digestion)

```
   1  MGAIGLLWLL PLLLSTAAVG SGMGTGQRAG SPAAGPPLQP REPLSYSRLQ RKSLAVDFVV
  61  PSLFRVYARD LLLPPSSSEL KAGRPEARGS LALDCAPLLR LLGPAPGVSW TAGSPAPAEA
 121  RTLSRVLKGG SVRKLRRAKQ LVLELGEEAI LEGCVGPPGE AAVGLLQFNL SELFSWWIRQ
 181  GEGRLRIRLM PEKKASEVGR EGRLSAAIRA SQPRLLFQIF GTGHSSLESP TNMPSPSPDY
 241  FTWNLTWIMK DSFPFLSHRS RYGLECSFDF PCELEYSPPL HDLRNQSWSW RRIPSEEASQ
 301  MDLLDGPGAE RSKEMPRGSF LLLNTSADSK HTILSPWMRS SSEHCTLAVS VHRHLQPSGR
 361  YIAQLLPHNE AAREILLMPT PGKHGWTVLQ GRIGRPDNPF RVALEYISSG NRSLSAVDFF
 421  ALKNCSEGTS PGSKMALQSS FTCWNGTVLQ LGQACDFHQD CAQGEDESQM CRKLPVGFYC
 481  NFEDGFCGWT QGTLSPHTPQ WQVRTLKDAR FQDHQDHALL LSTTDVPASE SATVTSATFP
 541  APIKSSPCEL RMSWLIRGVL RGNVSLVLVE NKTGKEQGRM VWHVAAYEGL SLWQWMVLPL
 601  LDVSDRFWLQ MVAWWGQGSR AIVAFDNISI SLDCYLTISG EDKILQNTAP KSRNLFERNP
 661  NKELKPGENS PRQTPIFDPT VHWLFTTCGA SGPHGPTQAQ CNNAYQNSNL SVEVGSEGPL
 721  KGIQIWKVPA TDTYSISGYG AAGGKGGKNT MMRSHGVSVL GIFNLEKDDM LYILVGQQGE
 781  DACPSTNQLI QKVCIGENNV IEEEIRVNRS VHEWAGGGGG GGGATYVFKM KDGVPVPLII
 841  AAGGGGRAYG AKTDTFHPER LENNSSVLGL NGNSGAAGGG GGWNDNTSLL WAGKSLQEGA
 901  TGGHSCPQAM KKWGWETRGG FGGGGGGCSS GGGGGGYIGG NAASNNDPEM DGEDGVSFIS
 961  PLGILYTPAL KVMEGHGEVN IKHYLNCSHC EVDECHMDPE SHKVICFCDH GTVLAEDGVS
1021  CIVSPTPEPH LPLSLILSVV TSALVAALVL AFSGIMIVYR RKHQELQAMQ MELQSPEYKL
1081  SKLRTSTIMT DYNPNYCFAG KTSSISDLKE VPRKNITLIR GLGHGAFGEV YEGQVSGMPN
1141  DPSPLQVAVK TLPEVCSEQD ELDFLMEALI ISKFNHQNIV RCIGVSLQSL PRFILLELMA
1201  GGDLKSFLRE TRPRPSQPSS LAMLDLLHVA RDIACGCQYL EENHFIHRDI AARNCLLTCP
1261  GPGRVAKIGD FGMARDIYRA SYYRKGGCAM LPVKWMPPEA FMEGIFTSKT DTWSFGVLLW
1321  EIFSLGYMPY PSKSNQEVLE FVTSGGRMDP PKNCPGPVYR IMTQCWQHQP EDRPNFAIIL
1381  ERIEYCTQDP DVINTALPIE YGPLVEEEEK VPVRPKDPEG VPPLLVSQQA KREEERSPAA
1441  PPPLPTTSSG KAAKKPTAAE ISVRVPRGPA VEGGHVNMAF SQSNPPSELH KVHGSRNKPT
1501  SLWNPTYGSW FTEKPTKKNN PIAKKEPHDR GNLGLEGSCT VPPNVATGRL PGASLLLEPS
1561  SLTANMKEVP LFRLRHFPCG NVNYGYQQQG LPLEAATAPG AGHYEDTILK SKNSMNQPGP
```

Curriculum Vitae

Ran Wei

Education

2010-2016 Ph.D Candidate, Biochemistry Department, Western University

2007-2010 MSc., Biochemistry Department, Western University

2001-2007 BSc/MSc., Biological Science Department, Wuhan University

Awards Received

1. 2014: Translation Breast Cancer Research Unit Scholarship

2. 2014: Poster Presentation Award, London Health Research Day

3. 2013: Translation Breast Cancer Research Unit Scholarship

4. 2012: Translation Breast Cancer Research Unit Scholarship,

Publications

1. Dhami, G. K., Liu, H., Galka, M., Voss, C., **Wei, R**., Muranko, K., Kaneko, T., Cregan, S. P., Li, L., and Li, S. S. (2013). Dynamic methylation of Numb by Set8 regulates its binding to p53 and apoptosis. Molecular cell 50, 565-576.

2. Liu, H. D., Galka, M., Mori, E., Liu, X. G., Lin, Y. F., **Wei, R**., Pittock, P., Voss, C., Dhami, G., Li, X., et al. (2013). A Method for Systematic Mapping of Protein Lysine Methylation Identifies Functions for HP1 beta in DNA Damage Response. Molecular cell 50, 723-735.

Publications in Preparation

1. **Wei, R**., Liu, X. G., Liu, H. D., Kaneko, T., Li, L., Voss, C., and Li, S. S. (in preparation). Investigation of NUMB-PTB interactome.

2. **Wei, R**., Liu, X. G., Liu, H. D., Voss, C., and Li, S. S. (in preparation). NUMB regulates the endocytosis and activation of Anaplastic Lymphoma Kinase in an isoform-dependent manner.