HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837

"ALEXANDER FLEMING"
Biomedical Sciences Research Center

NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

FACULTY OF MEDICINE

MEDICAL SCHOOL

MASTER'S PROGRAM in BASIC MEDICAL SCIENCES
"MOLECULAR BIOMEDICINE: MECHANISMS OF DISEASE, MOLECULAR AND CELLULAR THERAPIES, AND BIOINNOVATION"

MASTER'S THESIS

Studying the importance of a novel nuclear RNP (ribonucleoprotein complex) for tumor development and progression

NEFELI BONI-KAZANTZIDOU
Chemist, BSc

Supervisor: Dr PANAGIOTA KAFASLA
Researcher C
BSRC "Alexander Fleming"

ATHENS

SEPTEMBER 2018

Abstract

Alternative splicing (AS) is an important post-transcriptional regulatory process of gene expression, which has recently emerged as a "hallmark of cancer". Regulation of AS is coordinated by ribonucleoprotein complexes (RNPs) in response to changes in environmental or physiological conditions communicated through signaling pathways. However the mechanisms through which signaling pathways communicate with the splicing machinery and drive the interactions which orchestrate the switch of the AS profile to promote tumour growth, invasion and metastasis remain largely unknown.

This master's thesis project aimed to investigate a novel finding of this lab: the interaction of the scaffold protein Iqgap1 with one of the splicing regulators, hnRNPM, in the nucleus of cancer cell-lines. This finding has interesting implications, since the role of scaffold proteins in post-transcriptional regulation in the nucleus has not been explored. Based on information highlighting the importance of a switch in AS profile for gastric cancer patient survival, we applied co-immunoprecipitation and proteomics approaches to a normal gastric epithelial (HFE-145) and a gastric cancer cell line (NUGC-4) and identified 138 nuclear proteins interacting with IQGAP1, including 44 components of splicing machinery. Validation experiments confirmed the protein-protein interaction of IQGAP1 with hnRNPM, as well as RNA-dependent interactions with SRSF1, hnRNPA1, hnRNPA2B1, hnRNPC1/C2, DDX17 and other splicing regulators. Our results indicate that IQGAP1 participates in the formation of at least two different RNPs with potential roles in splicing and/or RNA metabolism. Phosphoproteomics analysis of IQGAP1 interacting partners also showed IQGAP1 interaction with phosphorylated spliceosome components, including PRPF6 and PRPF31.

In order to study the role of IQGAP1 depletion on gastric cancer cell phenotype and splicing patterns, CRISPR-Cas9 IQGAP1 knock-out clones of the cell lines of interest were subjected to phenotypical assays. The depletion of IQGAP1 appeared to have no effect on cell cycle progression in asynchronous culture, while, in contrast to HFE-145, the clonogenic capacity of NUGC-4 cells appeared to be increased after IQGAP1 depletion. At the same time, the results of a preliminary assay of hnRNPM-dependent splicing using the DUP51M splicing construct indicate that IQGAP1 depletion affects hnRNPM-dependent splicing partners.

Our current hypothesis is that IQGAP1 plays a significant role in the assembly of RNP complexes participating in pre-mRNA splicing and post-transcriptional regulation in gastric epithelial and gastric cancer cells, orchestrating patterns of AS which may be linked to phenotypical changes of gastric cancer cell lines. Interactions of IQGAP1 with post-transcriptional regulators emerge, therefore, as potential drivers of AS in gastric cancer- a result which, after systematic study, could contribute potential candidates for the development of novel therapeutics against a rare and deadly disease.

Περίληψη

«Μελέτη του ρόλου ενός νέου πυρηνικού ριβονουκλεοπρωτεϊνικού συμπλόκου για την δημιουργία και ανάπτυξη καρκινικών κυττάρων»

Το εναλλακτικό μάτισμα (ΕΜ) είναι μια σημαντική διεργασία μετα-μεταγραφικής ρύθμισης των γονιδίων, η οποία έχει προσφάτως αναδειχθεί ως κρίσιμη για τη δημιουργία καρκινικών κυττάρων. Η ρύθμιση του ΕΜ συντονίζεται από ριβονουκλεοπρωτεϊνικά σύμπλοκα (ΡΝΠ) αποκρινόμενα σε αλλαγές στις φυσιολογικές ή περιβαλλοντικές συνθήκες του οργανισμού ή ιστού που διαδίδονται μέσω σηματοδοτικών μονοπατιών. Ωστόσο, οι ακριβείς μηχανισμοί μέσω των οποίων τα σηματοδοτικά μονοπάτια επικοινωνούν με τα πρωτεϊνικά σύμπλοκα του ΕΜ και συντονίζουν τις πρωτεϊνικές αλληλεπιδράσεις οι οποίες ενορχηστρώνουν τις αλλαγές στο ΕΜ που υπεύθυνες για ανάπτυξη και μετάσταση των καρκινικών όγκων παραμένουν κατά βάση αδιευκρίνιστοι.

Ο σκοπός της παρούσας ερευνητικής εργασίας ήταν να διερευνήσει ένα νέο εύρημα του εργαστηρίου, την αλληλεπίδραση της κυτταροσκελετικής πρωτεΐνης IQGAP1 με μία από τις ρυθμιστικές πρωτεΐνες του ΕΜ, την ετερογενή πυρηνική ριβονουκλεοπρωτεΐνη M (heterogeneous nuclear ribonucleoprotein M, hnRNPM), στον πυρήνα καρκινικών κυτταρικών σειρών. Το εύρημα αυτό παρουσιάζει ενδιαφέρον, διότι ο ρόλος κυτταροσκελετικών πρωτεϊνών στη μετα-μεταγραφική ρύθμιση στον πυρήνα δεν έχει διευρενηθεί. Βασιζόμενοι σε πληροφορίες που αναδεικνύουν τη σημασία του ΕΜ στην καρκινογένεση του καρκίνου του στομάχου, εφαρμόσαμε τεχνικές ανοσοκατακρήμνισης πρωτεϊνών και πρωτεομικής ανάλυσης με φασματομετρία μάζας σε μια κυτταρική σειρά φυσιολογικού στομαχικού επιθηλίου (HFE-145) και μια κυτταρική σειρά προερχόμενη από ιστό λεμφαδένων στομαχικού αδενοκαρκινώματος (NUGC-4). Τα αποτελέσματα ανέδειξαν τις αλληλεπιδράσεις της IQGAP1 με 138 πυρηνικές πρωτεΐνες, μεταξύ αυτών και 44 γνωστές ρυθμιστικές πρωτεΐνες του ΕΜ. Περαιτέρω πειράματα επιβεβαίωσαν την πρωτεϊνική αλληλεπίδραση της IQGAP1 με την hnRNPM, καθώς και αλληλεπιδράσεις εξαρτώμενες από την παρουσία ριβονουκλεοπρωτεϊνικών μορίων (RNA) με τις ρυθμιστικές πρωτεΐνες SRSF1, hnRNPA1, hnRNPA2/B1, hnRNPC1/C2 και DDX17. Τα αποτελέσματα αυτής της μελέτης υποδεικνύουν ένα ρόλο για την IQGAP1 στο σχηματισμό ενός ή και περισσότερων νέων πυρηνικών ριβονουκλεοπρωτεϊνικών συμπλόκων με πιθανό ρόλο στη ρύθμιση του ΕΜ ή και του μεταβολισμού του RNA. Πρωτεομική ανάλυση με φασματομετρία μάζας των φωσφορυλιώσεων των πρωτεϊνών των ριβονουκλεοπρωτεϊνικών συμπλόκων της IQGAP1 έδειξε επίσης την αλληλεπίδρασή της με φωσφορυλιωμένες ρυθμιστικές πρωτεΐνες του ΕΜ.

Ο ρόλος της απόσβεσης της έκφρασης της IQGAP1, μέσω της τεχνολογίας CRISPR-Cas9, στον καρκινικό φαινότυπο και το μοτίβο του ΕΜ των μελετούμενων κυτταρικών σειρών μελετήθηκε με φαινοτυπικές δοκιμασίες. Η απόσβεση της έκφρασης της IQGAP1 δεν είχε αντίκτυπο στην εξέλιξη του κυτταρικού κύκλου σε ασύγχρονη καλλιέργεια. Η προκαταρκτική μελέτη της ικανότητας σχηματισμού αποικιών των τροποποιημένων κυττάρων είχε διαφορετικά αποτελέσματα μεταξύ των δύο κυτταρικών σειρών, ένα αποτέλεσμα που χρίζει περαιτέρω διερεύνησης. Παράλληλα, τα αποτελέσματα δοκιμασιών ΕΜ εξαρτώμενου από την hnRNPM ΕΜ χρησιμοποιώντας το πλασμίδιο DUP51M υποδεικνύουν ότι η απόσβεση έκφρασης της IQGAP1 επηρεάζει τα μοτίβα ΕΜ της hnRNPM.

Τα αποτελέσματα της παρούσας εργασίας υποδεικνύουν ότι η IQGAP1 έχει έναν σημαντικό ρόλο στο σχηματισμό πυρηνικών ριβονουκλεοπρωτεϊνικών συμπλόκων που συμμετέχουν στο ΕΜ και τη μετα-μεταγραφική ρύθμιση, ενορχηστρώνοντας μοτίβα ΕΜ που μπορεί να συνδέονται με φαινοτυπικές αλλαγές σε φυσιολογικές και καρκινικές κυτταρικές σειρές του στομάχου. Οι αλληλεπιδράσεις της IQGAP1 με ρυθμιστικές πρωτεΐνες του ΕΜ αναδεικνύονται, επομένως ως πιθανοί οδηγοί του ΕΜ στον καρκίνο του στομάχου, ένα αποτέλεσμα το οποίο θα μπορούσε να συμβάλλει στην εύρεση νέων στόχων φαρμακολογικής παρέμβασης ενάντια σε μια σπάνια και θανατηφόρο ασθένεια.

Acknowledgments

I would like to express my outmost appreciation and thanks to the supervisor of this work, Dr. Kafasla, for her unwavering and unconditional support and patience during the completion of this work. In addition, I would also like to thank Dr Martina Samiotaki for her generosity and valuable technical assistance, as well as Dr George Panayotou and Dr Thanos Kotsinas, members of my advisory committee. Finally, I would like to extend my thanks to my parents, for supporting me in this endeavor as they have in all others, and my friends, for their emotional support and patience.

Table of contents

Chapter 1. Introduction **15**
1.1 Alternative splicing 15
1.1.1 General mechanism 15
1.1.2 The spliceosomal proteome 17
1.2 Regulation of alternative splicing 20
1.2.1 General mechanism 20
1.2.2 Alternative splicing regulation through post-translational modifications 24
1.3 Alternative splicing and cancer 26
1.3.1 Hallmarks of cancer and alternative splicing 26
1.3.2 The role of heterogeneous nuclear ribonucleoprotein M (hnRNPM) in cancer pathogenesis 27
1.4 IQGAP1 30
1.4.1 Function and cellular localization of IQGAP1 30
1.4.2 IQGAP1 and cancer 32
1.4.3 Implications of IQGAP1's interaction with hnRNPM 33
Chapter 2. Aims of the project **35**
Chapter 3. Materials and methods **37**
3.1. Cell culture 37
3.2. Preparation of nuclear and cytoplasmic extracts 37
3.3. Immunoprecipitation of proteins 37
3.4. Western blotting 38
3.5. Mass spectrometry and Proteomics analysis 39
3.6. Analysis of phosphorylation in IQGAP1-immunoprecipitated proteins 40
3.7. Generation of cell lines with stable IQGAP1 knockdown by CRISPR-Cas9 41
3.8. Colony formation assay 45
3.9. Cell cycle analysis 45
3.10. Mini-gene splicing assay 46
3.11. RNA isolation/Reverse Transcription/PCR 46
Chapter 4. Results **48**
4.1 Study of IQGAP1-hnRNPM RBP complex 48
4.1.1 IQGAP1 enters the nucleus and interacts with hnRNPM 48
4.1.2 Nuclear IQGAP1 interacts with spliceosomal components 49
4.1.3 IQGAP1 forms two putative complexes with spliceosomal components 55
4.1.4 Study of post-translational modifications in IQGAP1-RBP complex components 57

4.2 Phenotypical study of IQGAP1 knockout cell lines 58
4.2.1 Generation of IQGAP1 knockout cell lines 58
4.2.2 Effects of IQGAP1 depletion on cell cycle progression on normal gastric epithelial and gastric cancer cells 59
4.2.3 Effects of IQGAP1 depletion on proliferative capacity of normal gastric epithelial and gastric cancer cells 60
4.2.4 Splicing effect of IQGAP1 knockout on hnRNPM-dependent splicing 61
Chapter 5. Discussion 63
Table of abbreviations 67
Addendum 69
Bibliography 71

Figure Index

Figure 1. Splicing by the major spliceosome 16
Figure 2. Biochemical mechanism of pre-mRNA splicing. 17
Figure 3 The two main groups of splicing regulatory proteins, heterogeneous nuclear ribonucleoproteins(hnRNPs) and serine/arginine rich splicing factors (SRSFs). 18
Figure 4. The spliceosomal proteome. 20
Figure 5. Mechanism of AS regulation... 21
Figure 6. Types of alternative splicing.. 23
Figure 7 : Signaling-activated kinases regulate splicing factor activity.. 24
Figure 8. Association of hnRNPM with human cancers. 28
Figure 9. Kaplan-Meier plot for hnRNPM mRNA expression in gastric cancer cases.. 29
Figure 10 Structure and functional domains of IQGAP1, including the main partners of each binding domain. 30
Figure 11 Association of IQGAP1 with human cancers. 33
Εικόνα 12. . Experimental procedure for proteomics analysis of IQGAP1 IPs 39
Figure 13. Principle of the employed CRISPR-Cas9 technique. 43
Figure 14. Sequence map of the empty vector All-in-One plasmid 44
Figure 15. IQGAP1 interacts with hnRNPM.. 48
Figure 16. IQGAP1 interacting proteins in HFE-145 (normal) nuclear extracts.. 49
Figure 17. IQGAP1 interacting proteins in NUGC-4 (cancer) nuclear extracts... 50
Figure 18. Quantitative comparison of IQGAP1 IP enriched proteins between HFE-145 (normal) and NUGC-4 (cancer). 51
Figure 19. Analysis of differentially enriched proteins in HFE-145 IQGAP1 IPs vs NUGC-4 IQGAP1 IPs. 52
Figure 20. Gene Ontology and KEGG pathway enrichment analysis for the 138 IQGAP1 nuclear interactors.. 54
Figure 21. Investigation of the composition of IQGAP1 nuclear RBP complex. 55
Figure 22. Generation of IQGAP1 knockout cell lines. 58
Figure 23. Cell cycle progression in IQGAP1 knockout cells. 59
Figure 24 Assessment of proliferative capacity of HQ and NQ cells. 60
Figure 25 Splicing effect of IQGAP1 depletion on hnRNPM-dependent splicing, using the DUP51 three exon, two intron splicing construct. 61
Figure 26. The proposed model for a nuclear role for IQGAP1 in splicing regulation.. 64

Table Index

Table 1. Modular enrichment analysis of 135 nuclear IQGAP1 interactors (for all gene ontology annotations) performed in Genecodis online platform 53

Table 2. Summary table of western blot validations of RNA-dependent or independent interactions as seen by co-immunoprecipitation experiments. 56

Chapter 1. Introduction

1.1Alternative splicing

1.1.1 General mechanism

Human cellular function exhibits a remarkable diversity, complexity and adaptability across cell types, tissues, developmental stages and disease states, such as cancer. This complexity cannot be explained solely by the diversity of the human genome. In 2001, at the time of publication of the first draft sequence of human genome, the number of protein-coding genes was estimated between 30.000-40.000 genes.[1,2] By 2004, when the Human Genome Project was declared complete, that estimate had fallen to 20.000-25.000 genes.[3,4] Over the past decade and a half, as the methods used to identify protein-coding genes have evolved, that number has decreased further. Today, the protein-coding sequences in the human genome are considered to be as few as 19.000-20.000 genes.[5,6]

Given this relatively limited number of protein-coding genes, the complexity of the human proteome, and that of cellular function, is ensured by mechanisms of genetic regulation at several levels. Multiple protein products, with potentially distinct functionality or localization may arise from a single protein-coding gene, as a result of regulation at the level of alternative transcription, splicing, 3' end formation, translation or post-translational modifications.[7,8] As the use of mass spectrometry based proteomics techniques [9] has continued to expand our knowledge of the human proteome[10–12], the actual number of protein isoforms and protein species (collectively known as proteoforms[13]) across tissues, cell types, and conditions has been difficult to ascertain, ranging from 70.000 to several million.[8] Post-transcriptional regulation, which includes alternative splicing (AS) has a major role in this variation: Between 92 and 95% of all protein coding genes have been found to undergo AS[14,15], and mRNA-seq data produced an estimate of ~100.000 alternative splicing isoforms with intermediate to major frequency in major human tissues.[14] Most importantly, this diversity of alternative pre-mRNA splicing events has been shown to have significant effects on the composition and function of the human proteome as well as cellular properties.[16–19]

AS is the first step of post-transcriptional regulation, which decides the fate of the gene transcript, known as precursor-messenger RNA, by determining which coding sequences (exons) will be included in the final mRNA sequence, and which non-coding sequences (introns) will be excised (reviewed in [20–22]). This decision is determined, in its first step, by the recognition of the splice sites.[23,24] These are conserved sequences at the 5' and 3' ends of intron-exon junctions which are the sites of assembly of the spliceosome, a macromolecular complex of nuclear RNA-binding proteins (RBPs) which is the core of RNA splicing machinery, responsible for the catalysis of the pre-mRNA splicing reaction.[25–28] The core of this complex is comprised of 5 small nuclear ribonucleoprotein subunits[29] (snRNPs) which interact with various other trans and cis acting factors. Each snRNP particle is formed through the interaction of uridine-rich small nuclear RNAs (U-snRNAs) with seven Sm or Sm-like proteins and other particle-specific factors.[30,31] The major spliceosome, which is responsible for the splicing of the majority of pre-mRNA introns (known as U2-type introns[32]) consists of 5 snRNPs (U1, U2, U4, U5, U6).[33–36] The minor spliceosome, responsible for the splicing of a rare type of introns (known as U12-type introns) is assembled by the less abundant snRNPs U11, U12, U4atac, U6atac and U5(reviewed in [32]).

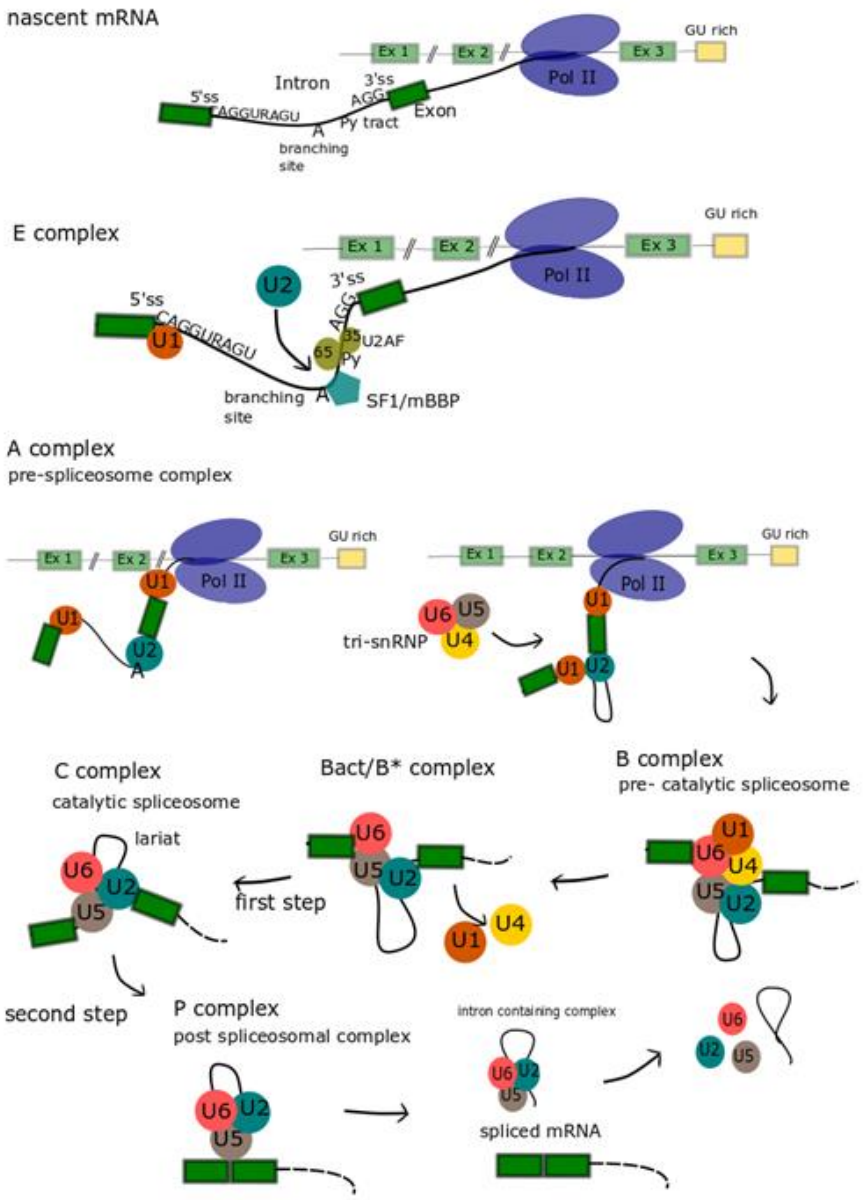

Figure 1. Splicing by the major spliceosome (see text).5' and 3' ss=splice site, Py tract= polypyrimidine tract, PolII= polymerase II, Ex1, 2, 3= exon 1, 2, and 3, respectively.

The catalysis of the splicing reaction by the major spliceosome has been extensively studied in vitro[37] and shown to be a step-wise process, with each step tightly coordinated by various splicing factors and RNA binding proteins.[20,38] About ~80% of splicing appears to occur co-transcriptionally[39], as the mRNA transcript is being released by RNA polymerase II, with the assembly of the spliceosomal components on the 5' and 3' splice sites of the nascent mRNA transcript. The 5' splice site is characterized by a highly conserved GU di-nucleotide at the intron end, while the 3' splice site includes three conserved sequence elements: the branching site, which contains a conserved adenine residue, followed by a polypyrimidine (Py) tract and a terminal AG dinucleotide.[24,40] (Figure 1). The splicing process begins with a Watson-Crick base-pairing interaction taking place between the U1 snRNA and the 5' splice site conserved sequence.[41] Subsequently, the 3' splice site is defined by the interaction of non-snRNP factors SF1/mBBP with the branch site sequence, while the 65kDa and 35kDa subunits of the U2AF splicing factor occupy the Py tract and the AG/G terminal dinucleotide, respectively.[42–46] This initial spliceosomal complex is known as the early complex (E complex), and its assembly, as with all subsequent complexes, is coordinated by protein-protein and protein RNA interactions of snRNAs, RBPs and splicing factors.[40,47] In a step which requires ATP

hydrolysis, the U2snRNP is recruited to the 3' splice site by the two U2AF subunits, displacing the SF1/mBBP and interacting with the branch site in an ATP dependent manner, thereby forming the pre-spliceosomal complex (complex A).[48,49] The pre-assembled U4/U5/U6 tri-snRNP complex binds to the 5'ss[50] to form the pre-catalytic B complex.[51–53] Several rearrangements take place in order to convert the spliceosomal complex to its catalytically active form, including the dissociation of the U1 and U4 snRNPs and the association of the PrP19/CDC5L complex, which enables the association of U5 and U6 snRNPs with the 5' splice site.[54,55] This leads to the formation of the catalytically active forms of the B complex, B* and B act complex(also known as intermediate C complex).[56–58]

This complex catalyzes the first of the two transesterification reaction steps which are necessary for intron excision[20,21] (Figure 2): In the first step, a nucleophilic attack takes place between the branch site adenine residue and the 5' splice site phosphate, producing a detached 5' exon and a 3' intron-exon intermediate in a lariat configuration (attached by the branch point adenine residue). After the lariat formation, the catalytically active complex C is formed[52], and proceeds with the second step of the splicing reaction. In this step, the 3' hydroxyl group of the now detached exon attacks the phosphate at the 3' of the intron, ligating the two exons together.[59–61] This leads to the formation of the post-spliceosomal complex (P complex), which releases the now mature mRNA and intron containing complex.[62] The process of 3' polyadenylation and the overall metabolic fate of the released mRNA is mediated by RBPs forming the exon-junction complex (EJC, reviewed in[63]) and the transcription-export complex (TREX), which is ultimately responsible for the nuclear export of the mRNA.[64]

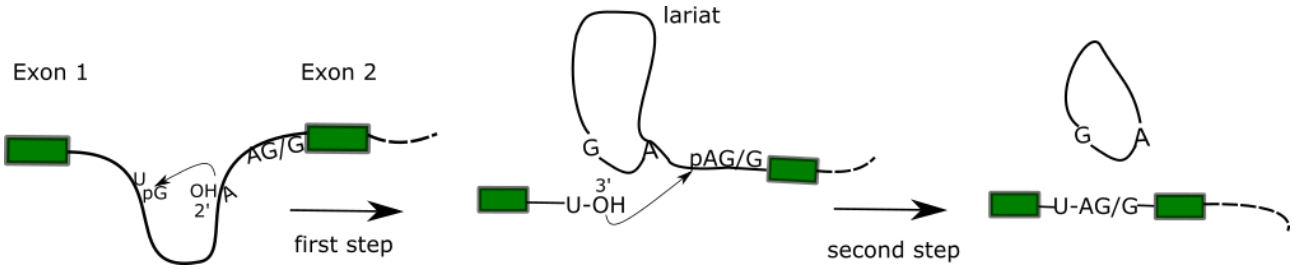

Figure 2. Biochemical mechanism of pre-mRNA splicing.In the first step, a nucleophilic attack by the 2' OH of the branch site adenine residue to the 5' splice site phosphate (of the conserved GU dinucleotide) forms the intron lariat, attached to the 3' exon and releases the 5' exon. In the second step the 3' OH of the detached exon attacks the phosphate group of the adenine residue in the conserved AG/G 3' sequence of the intron. The intron lariat is excised, and the two exons are ligated.

1.1.2 The spliceosomal proteome

The numerous proteins which interact with the core components of the spliceosome and participate in spliceosomal assembly constitute the spliceosomal proteome, which has been extensively studied in order to elucidate both the composition and mechanistic functions of each spliceosomal complex (reviewed in[27,38,65]). As RBPs, spliceosomal proteins exhibit similarities in structure and function[66,67], however they can be separated into several functional categories, which include:

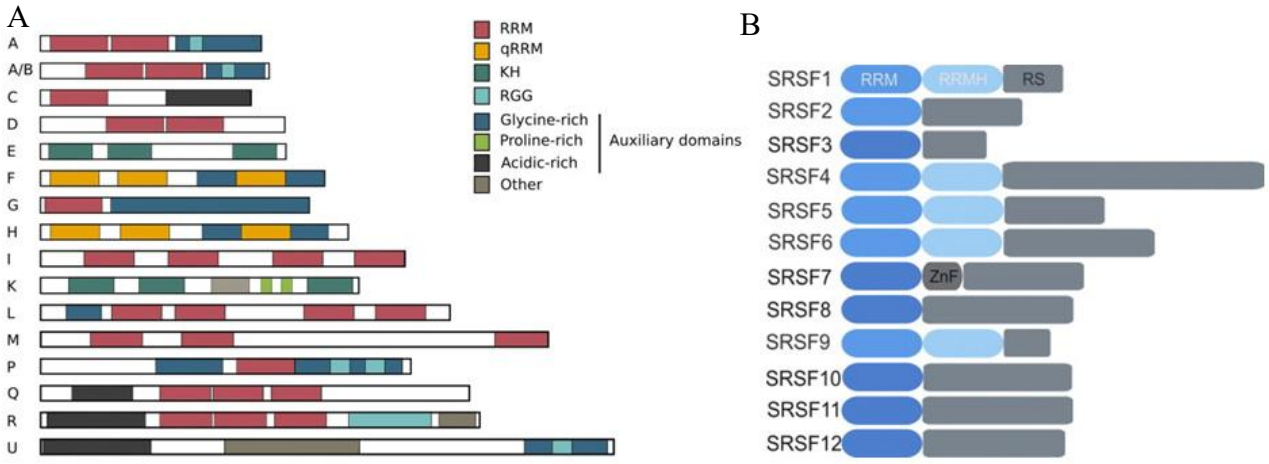

Figure 3 The two main groups of splicing regulatory proteins, heterogeneous nuclear ribonucleoproteins(hnRNPs) and serine/arginine rich splicing factors (SRSFs). A. The hnRNP family. The 20 main members of thehnRNP family are named alphabetically from hnRNP A1 to hnRNPU.. hnRNPs are characterized by four unique RNA-binding domains (RBDs):: RRM (RNA recognition motif), qRRM (quasi-RNA recognition motif), KH (K-homology domain), RGG (RNA-binding domain consisting of Arg-Gly-Gly repeats). The sizes of these 16 common hnRNPs are drawn relative to each other (adapted from[68].) B. The main members of the SRSF family and their functional domains: N-terminal RRM domains and C-terminal serine/arginine rich domains (RS domains).SRSF7 can also be equipped with a zinc finger (ZnF) domain, possibly involved in RNA-binding (adapted from[81]).

1. **Heterogeneous nuclear ribonucleoproteins (hnRNPs)**: These are a large family of abundant ribonucleoproteins, which perform a variety of molecular functions, including constitutive and alternative splicing, transcriptional and translational regulation and mRNA stability (reviewed in[68,69]). The first isolated hnRNPs (hnRNPA/B and C), were initially isolated as components of the the 40S core particle, by sucrose density gradients.[70] Subsequent pull-down and UV crosslinking experiments and identification using monoclonal antibodies allowed the isolation of increasing number of hnRNP complexes leading to the identification of 20 major and several minor types of hnRNPs.[71,7273] The major hnRNPs are alphabetically named hnRNPA-U, although many have alternative names, and their molecular weights range from 34-120 kDa (Figure 3A). The function of hnRNPs varies, depending on their functional domains and their cellular localization. The primary location of most hnRNP proteins in steady state is the nucleus, and most of them are able to shuttle in and out of the nuclear envelope thanks to the presence of a conventional nuclear localization signal (NLS), as in the case of hnRNPK, or via transcriptionally dependent translocation (as in the case of hnRNPA1).[69] Translocation of hnRNPs to the cytosol can occur upon post-translational stimulation or by recruitment of other hnRNPs. [68,69] Most hnRNPs have been shown to have important roles in splicing[74] as well as alternative splicing regulation (reviewed in[75]): thousands of hnRNP-dependent alternative splicing events have been identified, in which hnRNPs can act both as activators or suppressors of exons to which they bind[76,77] The interactions of hnRNPs with RNA are mediated through their RNA binding domains (RBDs): The most prevalent of these is the RNA Recognition Motif (RRM).[78] hnRNPs which do not have RRMs mediate their interactions via quasi/atypical RRM domains (which lack characteristic RNP consensus sequences or other RRM elements), K homology (KH) domains or RGG domains (consisting of Arg-Gly-Gly repeats).[68] Other secondary RBDs are also present in hnRNPs.
2. **Serine-Arginine Rich Splicing factors (SRSFs)**: SRSFs are a conserved family of RBPs, characterized by sequences of consecutive serine and arginine repeats.[79] The main SRSFs are seven phosphoproteins ranging from 20 to 75 kDa, which were initially identified thanks to the development of a monoclonal antibody (mAb104)[80] specific for their conserved phosphorylated serine-arginine rich sequence. Each of the SR proteins bears one or two RRMs and an arginine-

serine rich (RS) domain, which enables them to interact with each-other (Figure 3B). Like hnRNPs, SRSFs exhibit a variety of molecular functions, including transcriptional and translational regulation (reviewed in[81]). SRSFs have an indispensable role in splicing and alternative splicing regulation, by mediating the recruitment of the U2AF complex on the 3'splice site [45] and coordinating early complex formation.[40,46] The regulatory role of SRSFs is complex: In addition to promoting splicing, SR proteins can also suppress it, depending on their binding site within a given mRNA. [82,83] SRSF interactions with the pre-mRNA are mediated through their RRM domains.[84] The RS domain is highly phosphorylated and mediates localization and protein-protein interactions, however it is not required for many splicing functions.[85,86] Recent evidence suggests that SRSF interactions with 5' splice sites are not solely determined by the primary sequence of the protein-coding region of the pre-mRNAs, but are also dependent on SRSF interactions with single-stranded exonic RNA elements immediately upstream of the 5'ss, which modulate the pre-mRNA structure regardless of nucleotide sequence.[87]

3. **RNA helicases**: Another group of spliceosomal RBPs necessary for spliceosomal assembly are the RNA helicases (reviewed in[88]). These are a highly conserved family of enzymes that are able to bind and remodel RNA or RNPs through ATP hydrolysis. They mediate the alterations in RNA-RNA, RNA-protein and RNP conformational changes which need to occur in several steps of the spliceosome assembly.[21,38,65,89,90] Eukaryotic RNA helicases are closely related to the DNA helicases, and some helicases, such as DHX9, work on both nucleic acids.[88] Depending on their characteristic sequence motifs, RNA helicases are split into 6 families: Ski2-like, DEAH/RHA,NS3/NPH-II, DEAD-box, Upf1-like and retinoic-acid-inducible gene I (RIG-I)-like.[88] Most RNA helicases participate in various cellular processes, such as ribosome biogenesis, pre-mRNA splicing, and translation.[88] Eight evolutionarily conserved members of the DEAD-box, DEAH-RHA and Ski-2 like families have been identified as core components of spliceosomal assembly, including Sub2 (UAP56 in humans), Prp5, Prp28, Brr2, Prp2, Prp16, Prp22, Prp43 and the GTPase Snu114[38,65,90] Although initially identified in *S. cerevisiae*, these 8 RNA helicases are required for splicing in all eukaryotes, including *Homo sapiens*.[90]

4. **Other splicing factors**: Many other proteins are involved in spliceosomal assembly, including, of course, the snRNP-associated proteins. These include the seven Sm proteins (B/B′ , D1, D2, D3, E, F and G),as well as U snRNP-specific proteins. For example, the proteins required for U1snRNP formation include the RRM-containing U1A and U1-70K and the zinc-finger domain containing U1C.[65] The two subunits of the U2AF complex (U2AF[35] and U2AF[65]) are associated with the U2 snRNP complex, as well as the SF3a and SF3b particles which are required for formation of the active form of the 17S U2 snRNP.[65]

Each spliceosomal complex is formed as a result of tightly coordinated interactions between numerous proteins. Aside from U2-associated proteins and members of the hnRNP and SRSF families are present, in varying abundance, in all steps of the splicing cycle, a large number of additional protein components are recruited in the transitions between each spliceosomal complex (A to B to C), many of which are released or de-stabilized in subsequent steps, in a dynamic remodeling of complexes.[38]

The actual number of proteins which participate in the dynamic process of spliceosome assembly has been studied *in vitro* and *in vivo*, primarily with mass spectrometry based techniques[52,56,57,62,91–96], and may be as high as several hundred. In addition to mass-spectrometry based studies, a plethora of structural studies of *in vitro* assembled pre-catalytic and catalytic spliceosomal complexes which have been conducted with cryo-electron microscopy[40,53,58,97–101] and single molecule fluorescence microscopy-based techniques (reviewed in[102]) has also shed light on the mechanism and interactions involved in complex assembly in atomic

resolution. Thanks to these studies, the structure and function of the splicing machinery is being elucidated in detail, identifying an increasing number of protein components in each step of spliceosome assembly (Figure 4).

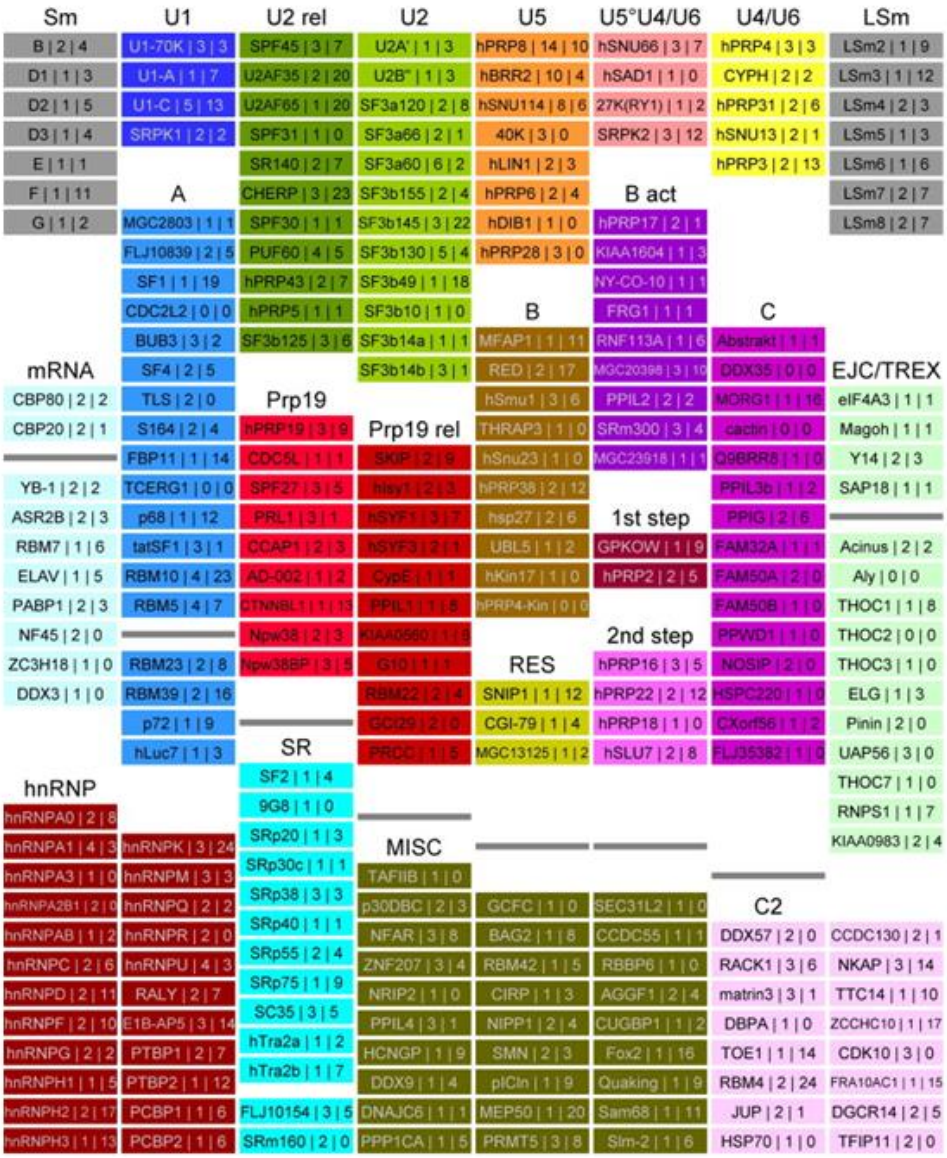

Figure 4. The spliceosomal proteome (adapted from[94]). Human proteins (244) that copurify with defined spliceosomal complexes are named according to the commonly used nomenclature[38] and grouped (each group color-coded) according to their presence in a given complex or their function. They are also classified into core and noncore proteins (above and below the gray bar, respectively) primarily based on their abundance in spliceosomal complexes (based on[103]). The pipes (j) separate the number of clones used and the number of PPIs found with the protein in theY2H matrix screen.

1.2 Regulation of alternative splicing

1.2.1 General mechanism

Given the number of players involved, it becomes obvious that regulation of alternative splicing is a multi-factorial process which hinges on the tight spatio-temporal coordination of a very large number of RNA-RNA, RNA-protein and protein-protein interactions. This is a huge and complicated task. The first crucial decision in the heart of alternative splicing control is the definition of exons and introns which are to be included in the final transcript. As mentioned, this is a decision mechanistically mediated by the recognition

and binding of the 5' and 3' splice-sites. However, this is not a straightforward process: Splice-sites are not solely sufficient for the definition of intron-exon borders[104]-instead, AS regulation is a complex, combinatorial process influenced by several factors (reviewed by[104–106]). Exon/intron definition is aided by the presence of auxiliary cis-acting consensus sequences on the exon-intron borders, proximal to 5' and 3'ss, binding of which either enhances or suppresses (silences) exon/intron inclusion. These are designated as exonic splicing enhancer and exonic splicing silencer sequences (ESE and ESS, respectively) as well as intronic splicing enhancer and intronic splicing silencer sequences (ISE and ISS), depending on their exon-intron location and activity. Silencers and enhancers are present both on constitutively and alternatively spliced exons/introns. ESEs act by assisting E complex formation through their interaction with the RRM domain of SRSFs, which subsequently participate in recruitment of the U2AF2 complex, as previously mentioned. The mechanism of function of ISEs has not been as well characterized, however they are thought to function in a similar manner, recruiting key splicing components to the 5' and 3' ss. On the other hand, the suppressing function of ESS and ISS sequences is thought to be mediated by hnRNPs, which also interact with them through their respective RRM or other RBP domains. Exactly how hnRNPs can perform their suppressive function is not yet fully clear, however evidence so far suggests that they generally suppress E complex formation and thus exon/intron inclusion by causing sterical hindrance, or the formation of inhibiting RNA secondary structures.[104] The regulatory role of hnRNPs and SRSFs however is not so one-sided - as we have mentioned, there are members of both protein groups which can act as exon promoters or suppressors, depending on their binding site and mode of action.[76,77,82,83]

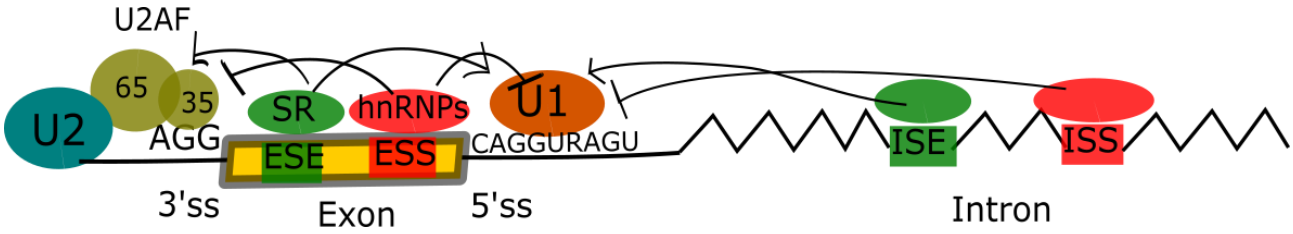

Figure 5. Mechanism of AS regulation. Binding of SRSFs and hnRNPs to cis-acting auxiliary regulatory sequences, namely exonic or intronic enhancers (ESE, ISE) or silencers (ESS, ISS), either promotes or inhibits association of the U1 snRNP to the conserved 5'ss sequence, or the recruitment of the U2 snRNP to the 3'ss by the U2AF particle. SRSFs are thought to primarily act as splicing enhancers (binding to ESE's) and the RRM-domain binding of hnRNPs to ESS and ISS mediates their activity as splicing silencers.

To summarize, enhancer and silencer sequences play a crucial role in alternative splicing regulation by recruiting RBPs which can promote the formation of the E complex and the initiation of splicing on their adjacent 5' and 3' splice sites (SRSFs) or, conversely, inhibit E complex assembly and thus lead to exclusion of the exon or intron defined by these borders from the final transcript (hnRNPs) (Figure 5). Aside from the presence of regulating sequences (5' and 3' splice sites and enhancer/silencer sequences) alternative splicing can also be influenced by the following factors:

a) the pre-mRNA secondary structure, which can act either by modulating the availability of splicing sequences or through the action of riboswitches, natural RNA aptamers which can affect splicing by regulating 3' UTR processing events[24]

b) priming events, ie the presence or absence of preceding splicing events, which may affect whether or not the alternative intronic sequences will be recognized[24]

c) coupling between transcription and alternative splicing outcome[24] and

d) effects of chromatin structure and histone modifications on spliceosomal complex assembly.[47,107,108]

Finally, there is emerging evidence of a higher order of regulation, indicating that the function of splicing regulators themselves can also influenced by AS of their own transcripts, and the presence of regulatory

motifs in their protein sequence: A recent study observed an enrichment of AS events in low complexity intrinsically-disordered regions (IDRs) of mammalian proteins, including several members of the hnRNP family.[109] It was demonstrated that this mode of regulation was mediated through the formation of long-range RNA duplexes, and resulted in the differential inclusion of exons rich in glycine (G) and tyrosine (Y) rich motifs- the presence of absence of these GY motifs in the final mRNA transcript, in turn, affected the ability of the expressed hnRNPs to form multivalent complexes, influencing downstream AS.[109]
It becomes apparent that alternative splicing regulation depends on a variety of factors which determine and are determined by the composition and functions of the spliceosomal proteome. The resulting decision of whether or not a given exonic or intronic sequence will be retained in the final transcript or excised leads to 5 basic types of alternative splicing events (Figure 6):

a) Exon skipping or cassette exon (Figure 6a): this has been observed be the most common type of alternative splicing in mammals.[110]In this type of AS event, an exon which lies between two intronic sequences may be included in the mature mRNA transcript or excised (as a result, different exons containing start-codon or stop codon sequences may be retained as the first or last exon, respectively).
b) Mutually exclusive exons (Figure 6b): In this subtype of AS, two or more splicing events are no longer independent, but are enhanced or suppressed in a coordinated manner[111]: For example, if one of the two mutually exclusive exons is included into the final transcript, the other will be spliced out of it, and vice-versa.
c) Alternative intron retention: In this type of AS event, an intron may be included or excised from the final transcript.
d) Alternative 5' splice site selection: intron excision may start at different points in the intron sequence.
e) Alternative 3' splice site selection: exon retention may start at different points in the exon sequence, variably including whole or part of the exon in the mature transcript.

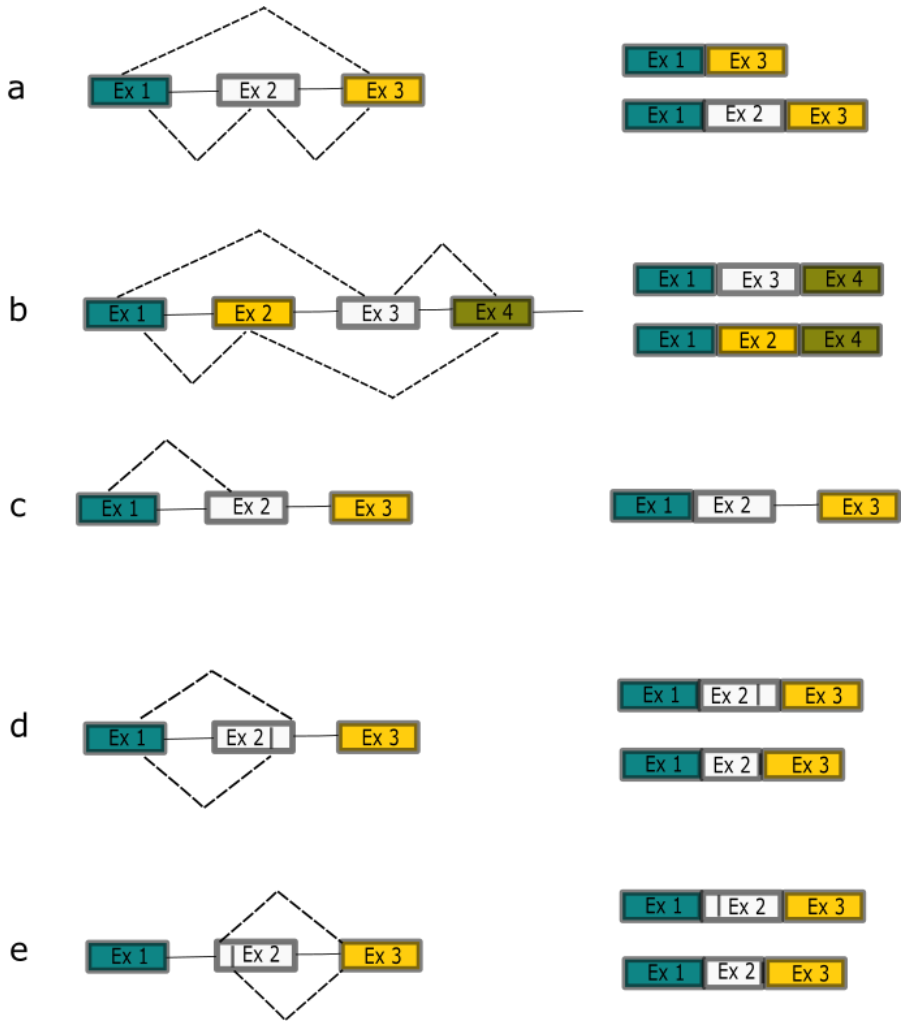

Figure 6. Types of alternative splicing. The perforated line indicates a splicing reaction between the two ligated exons. The products of the possible AS reactions (indicated by lines above and below the pre- mRNA) are shown on the second column (above and below respectively).a. Exon-skipping or cassette exon. b. Mutually exclusive exons. c. Alternative intron retention. d. Alternative 5' splice site selection. e. Alternative 3' splice site selection.

1.2.2 Alternative splicing regulation through post-translational modifications

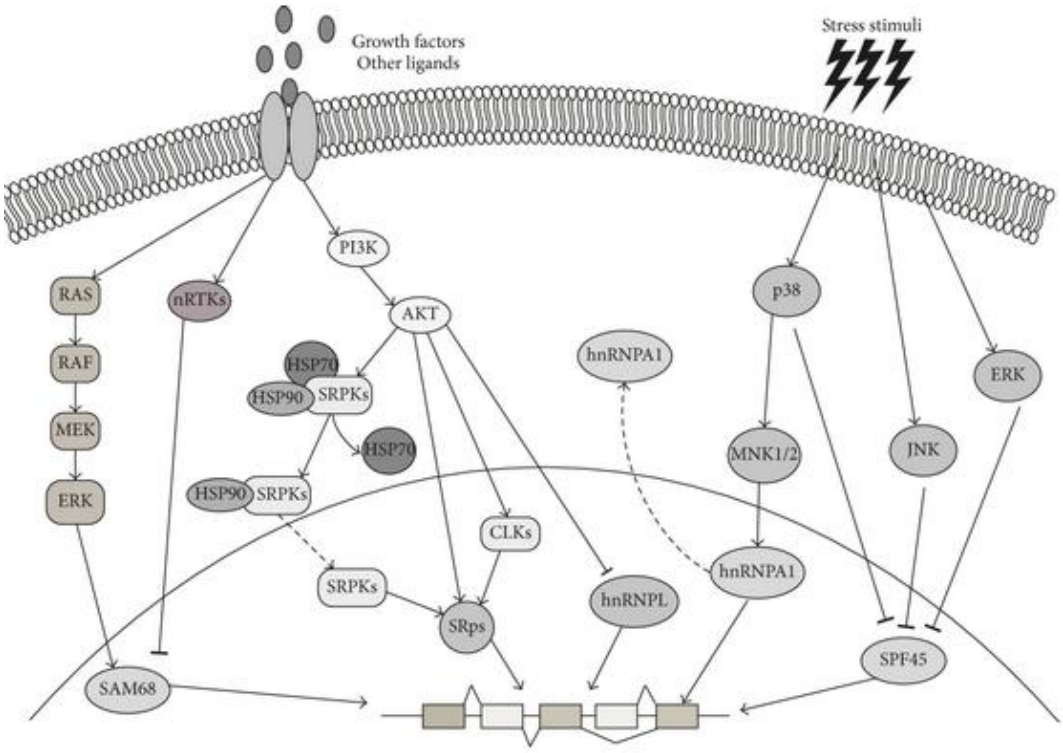

Figure 7 : Signaling-activated kinases regulate splicing factor activity (adapted from[115]). Various extracellular cues, like growth factors or stress stimuli, activate different signal-transduction cascades impinging on protein kinases that in turn phosphorylate RBPs, thereby modulating their splicing activity. SAM68 splicing activity is inversely regulated by ERKs and nRTKs, which, respectively, activate and inhibit its splicing activity. The PI3-K-AKT pathway regulates the activity of several SR proteins both directly or by phosphorylating and modulating the activity and localization of CLKs and SRPKs. Stress signal-activated kinases, like JNK or p38, can both modulate splicing factor localization (e.g. hnRNPA1), or activity(e.g. SPF45).

Of course, in order for the cell to be able to choose which AS events will be implemented in response to any given condition, the precise regulation of AS needs to be able to respond and adapt to organism, tissue and cell-wide events. This communication between changes in cellular or environmental conditions and alternative splicing programs is achieved through post-translational modifications of splicing factors mediated through signaling pathways (reviewed by[112–115]). Post-translational modifications (including phosphorylation, acetylation, hydroxylation and ubiquitination) act as "switches" which can control the stability and degradation of splicing proteins as well as their intracellular localization and protein-protein or protein-mRNA interactions.[114]

Phosphorylation is one of the key signaling events responsible for the modulation of the function of splicing proteins.[115–118] One of the most common post-translational modifications in splicing regulation is the phosphorylation of the serine residues of the RS domains of the SRSF proteins which is an effective way of modulating their localization as well as their protein-protein and protein RNA interactions.[118–120] For example, SRSF1's interaction with U2AF65 is enhanced by phosphorylation of its RS domain.[121,122] Phosphorylation of SRSFs is mediated by two main families of protein kinases, the SR-protein kinases (SRPKs) and cyclin-dependent- like kinases (CLKs).[115] Phosphorylation of SRSF1 by the kinase SRPK1, has been shown to regulate alternative splicing of Rho-GTPase Rac1.[123] SRPK-mediated SRSF1 phosphorylation has been shown to be crucial for the formation of nuclear speckles[124], while hypo-phosphorylation of SR

proteins promotes TAP-mediated nuclear export of spliced mRNA.[125] Function and subcellular localization of hnRNPs is also mediated by phosphorylation.[115] A prominent example is the subcellular localization of hnRNPA1, which changes in response to stress mediated signaling: hnRNPA1 phosphorylation reduces its interaction with its nuclear transporter, leading the nucleo-cytoplasmic shuttling protein to accumulate in stress granules in the cytoplasm, and affecting hnRNPA1-mediated AS .[126–128]

Phosphorylation of non-SR components is also required in several steps of the spliceosomal assembly: The U2 snRNP protein SF3b1/SAP155 is phosphorylated concomitantly with splicing catalysis[129,130] and phosphorylation of the U4/U5/U6 tri-snRNP complex components PRPF6 and PRPF31 by the PRP4 kinase is required for B complex assembly.[131] Another example is CDC5L, which undergoes cycle-dependent phosphorylation required for CDC5L-mediated mRNA splicing.[132] Phosphorylation events also play a role in the co-ordination of transcriptionally coupled splicing. Phosphorylation of the C-terminal of Pol II is required for its association with initial splicing components in co-transcriptional splicing.[133] Numerous RNA-binding proteins, including SR proteins and Sam68 interact with the Pol II complex or transcription factors to facilitate their recruitment to nascent RNA, in a phosphorylation dependent manner.[134–136]

Phosphorylation-dependent regulation of splicing factor function can be mediated through all known signaling pathways (Figure 7): Kinases of the mitogen activated protein kinase (MAPK) pathway, as well as the PI3K activated-Akt kinases have been shown to mediate Ser/Thr phosphorylation of SRSFs, hnRNPs and other splicing factors.[115] One example of Akt-dependent AS regulation is the growth signal response mediated by SRSF phosphorylation through the PI3k-Akt-SRPK axis.[137] The MAPK family of kinases includes the exracellular regulated kinases (ERK1/2), as well as the c-Jun amino terminal kinases (JNK 1-3), p38 and the ERK5 sub-family.[138] A classic example of MAPK-mediated regulation of AS involves the regulation of the CD44 gene, which includes 10 variable exons. The inclusion of variable exon v5 in the mature CD44 mRNA was shown to be dependent on the activation of Sam68 by the RAS-RAF-MEK-ERK signaling cascade.[139,140] Another example of MAPK signaling inducing splicing factor modifications is the phosphorylation of SPF45, which can be mediated by all three families of MAPKs in response to different stress stimuli, modulating SPF45-dependent AS of the FAS gene.[141] Members of the STAR family of RBPs, including Sam68, are among the few known to undergo tyrosine phosphorylation by non-receptor tyrosine kinases (nRTKs) members of the protein tyrosine kinase (PTK) family, while phosphorylation by the cAMP-dependent protein kinase (PKA) which responds to the levels of cyclic adenosine 3'-5' -monophosphate (cAMP), have been shown to modulate the function of both hnRNPs and SRSFs.[115]

Aside from phosphorylation, other post-translational modifications also play a role in modulating the activity of splicing proteins: Arginine methylation of SRSF1 controls its subcellular localization[142], while EGF-mediated ubiquitylation of hnRNPA1 has been shown to induce an alternative splicing program which leads to Rac1 upregulation, promoting cell motility.[143] Another example is the finding that the 65 kDa subunit of U2AF undergoes post-translation lysyl-5-hydroxylation mediated by 2-oxoglutarate-dependent dioxygenase Jumonji domain-6 protein (Jmjd6), which was shown to drive alternative splicing programs.[144]

To summarize, signal transduction is known to induce modifications of AS regulating proteins, leading to changes in AS regulation in response to MAPK pathway activation, heat-shock, depolarization, stress signals, or T cell signaling (Figure 7).[106,114] It is therefore evident that signal transduction pathways have a very important role in AS regulation. However, the specific mechanisms through which signaling events are integrated into system-wide changes in splicing regulators, which in turn induces global changes in RNP assembly and AS, are not yet fully understood.[106,114]This, therefore, is the first question that arises from the state of the current research: Which are the missing links which mediate the differential modification and assembly of splicing factors in response to a specific stimulus, bridging the gap between signaling changes and AS regulation?

1.3 Alternative splicing and cancer

1.3.1 Hallmarks of cancer and alternative splicing

Considering complex and delicate balance of AS coordination, and the consequences it can have in creating alternative protein isoforms or affecting transcript stability and localization, it is not surprising that dysregulations in the splicing process have been implicated in a variety of human diseases (reviewed in[145–148]).
This dysregulation can take many forms: the most common of these is mutations in the affected pre-mRNA sequence.[145] Cis-acting mutations in the consensus sequences, including the 5' and 3' splice sites and the branch point, as well as alterations in the regulatory regions responsible for spliceosome recruitment, -including exonic and intronic splicing enhancers and silencers- can lead to the generation of alternative transcripts causal to a single or multiple pathological phenotypes. An example of this is the cases of mutations in the dystrophin splice site which can lead to generation of aberrant transcripts with loss of dystrophin function, and development of Duchenne's muscular dystrophy.[149,150]

However, because of the great significance of protein-protein and protein-RNA interactions in the assembly of the spliceosome and alternative splicing regulation, mutations, alterations in expression and post-translational modifications of RNA binding proteins which affect such interactions have also been implicated in human pathology. Perturbations in the function of the ubiquitous and multifunctional components of the core spliceosome machinery, including heterogeneous nuclear ribonucleoproteins and splicing factors can lead to severe perturbations on cellular function.[148]

The role of alternative splicing mis-regulation in oncogenesis and cancer development and progression has been thoroughly demonstrated in recent years[151] (reviewed in[152–161]). Indeed, in 2014, Oltean and Bates stipulated that alternative splicing dysregulation represents an underlying mechanism which can drive all the processes which Hannahan and Weinberg originally named "the hallmarks of cancer", and which are required for tumor development and progression: enabling replicative immortality, sustaining proliferative signaling, evading growth suppressors and apoptosis, inducing angiogenesis, activating invasion and metastasis, deregulating cellular energetics and avoiding immune destruction. [156,162] This relationship between alternative splicing and cancer goes both ways, with oncogenic pathways in turn regulating alternative splicing to their advantage. [156]

As a result, alternative splicing has emerged a target of therapeutic intervention in cancer (reviewed by[157,163–166]). For example, a novel indole alkaloid, jerantinine A, was recently shown to induce tumor-specific cell death through modulation of SF3B1 activity.[167], while H3B-8800, a small-molecule splicing modulator, was shown to induce lethality in SF3b1-mutant drug resistant tumor cells.[168] AS regulation is also emerging as an effector of drug resistance- for example, TRA2A- mediated AS promotes paclitaxel resistance and tumor progression in triple-negative breast cancer.[169]

SRSF1 is one of the key spliceosomal components that has strongly been implicated in carcinogenesis, and is now considered an oncogene.[170–172] SRSF-driven AS events affect most hallmarks of cancer[156]. Many other splicing factors have been associated with cancer, including PRPF6[173], hnRNPA1[174], hnRNPA2/B1[175], hnRNPM and hnRNPL.[176–178]

1.3.2 The role of heterogeneous nuclear ribonucleoprotein M (hnRNPM) in cancer pathogenesis

One of the hnRNPs which has been recently associated with cancer pathogenesis is hnRNPM. HnRNPM is a family of human ribonucleoproteins which includes two distinct subsets: The M1-M4 polypeptides (a cluster of 4 proteins between 64 and 68 kDa)[179] and the 72/74 kDa polypeptides.[180–182] hnRNPM polypeptides have been shown to interact with pre-mRNA in spliceosomal formation, as well as being a part of the nuclear matrix.[183] As a member of the hnRNP family, hnRNPM is one of the most abundant spliceosomal proteins.[69] While hnRNPs generally exhibit multiple functional roles in transcriptional regulation and mRNA stability, hnRNPM has been primarily associated with splicing and AS regulation.[68,69]

In Drosophila, hnRNPM is involved in the alternative splicing of its own mRNA[184], as well as affecting the splicing of fibroblast growth receptor 2 (FGFR2)[185] and CD44[186] and enforcing alternative splicing programs in response to extracellular signaling.[187] hnRNPM can act as both an enhancer and a silencer of alternatively spliced exons[185], and affects both 5' and 3' splice site choices.[188]

More signficantly, hnRNPM has been shown to be implicated in tumor pathogenesis and progression in various cancer types, both *in vivo* and *in vitro*: hnRNPM appears to potentiate TGF β -mediated epithelial-to-mesenchymal transition in breast cancer, by promoting an alternative splicing isoform of CD44 in a mesenchymal-specific context.[186] In the same study, it was demonstrated that hnRNPM over-expression positively correlates with tumor aggressiveness in breast cancer patients. Indeed, hnRNPM and CD44 levels were subsequently shown to predict poor prognosis in breast cancer with axillary node metastases.[189] In addition, the M4 isoform of hnRNPM was identified as a receptor of the carcinoembryonic antigen (CEA), a regulator of intercellular adhesion which is a prognostic tumor marker in colorectal cancer, and was later proposed to promote colorectal cancer by targeting E-cadherin adherens junction complexes[190–192]. Significantly, a quantitative proteomics study of colorectal carcinoma patients identified hnRNPM as a tumor biomarker for colorectal cancer.[193] hnRNPM was also observed to enforce an AS program in response to inhibition of the mTOR/AKT/PI3K pathway in Ewing Sarcoma cells.[187] It was also shown to be correlated with cell-cycle and apoptosis regulators in a retrospective tissue microarray study, indicating a role in ovarian cancer pathogenesis.[194]

The impact of hnRNPM on cancer pathogenesis can also be examined by reviewing data from large harmonized data repositories of human cancer cases, including the Genomic Data Commons Portal, which includes datasets from 40 different projects, including the "Cancer Genome Atlas" and integrates profiles for 32.555 cancer cases corresponding to 61 different primary tumor sites.[195]

The GDC data on hnRNPM lists 189 cancer cases affected by 217 hnRNPM mutations across 25 cancer projects. The cancer type most highly affected by hnRNPM mutations is uterine corpus endometrial adenocarcinoma, with 53 out of 530 cases affected (10%), while stomach adenocarcinoma is the second most affected cancer, in 5.45% of cases (Figure 8A). Pathology data from the Human Protein Atlas[196–198] show that hnRNPM mRNA is expressed in all cancer types, most of which exhibit strong staining for hnRNPM protein expression. Expression of hnRNPM mRNA is an unfavourable prognostic marker in liver cancer (p value= 1.8 x 10-6, 365 samples), while it appears to be a weak favourable correlation in gastric cancer (stomach adenocarcinoma) (non-prognostic) (Figure 8B). Interestingly, Kaplan-Meier plots of patient survival based on transcriptomic data from cancer patient cases also shows a favourable prognostic value (log-rank p value = 5.2×10^{-11}) for hnRNPM mRNA expression in gastric cancer,[199–201] which is one of the three cancer types noted to be most affected by alternative splicing changes[161] (Figure 9). This might be a first indication that

hnRNPM has a protective role in gastric cancer pathogenesis, and that mutations or perturbations of hnRNPM splicing function might have a tumor promoting effect in gastric cancer.
In the face of this evidence, hnRNPM can be seen as a possible orchestrator of cancer-promoting AS events in various cancer types. A possible mechanism behind the connection of hnRNPM-mediated AS regulation and tumor development and progression emerges, therefore, as a point of interest, and formed the second part of our research question.

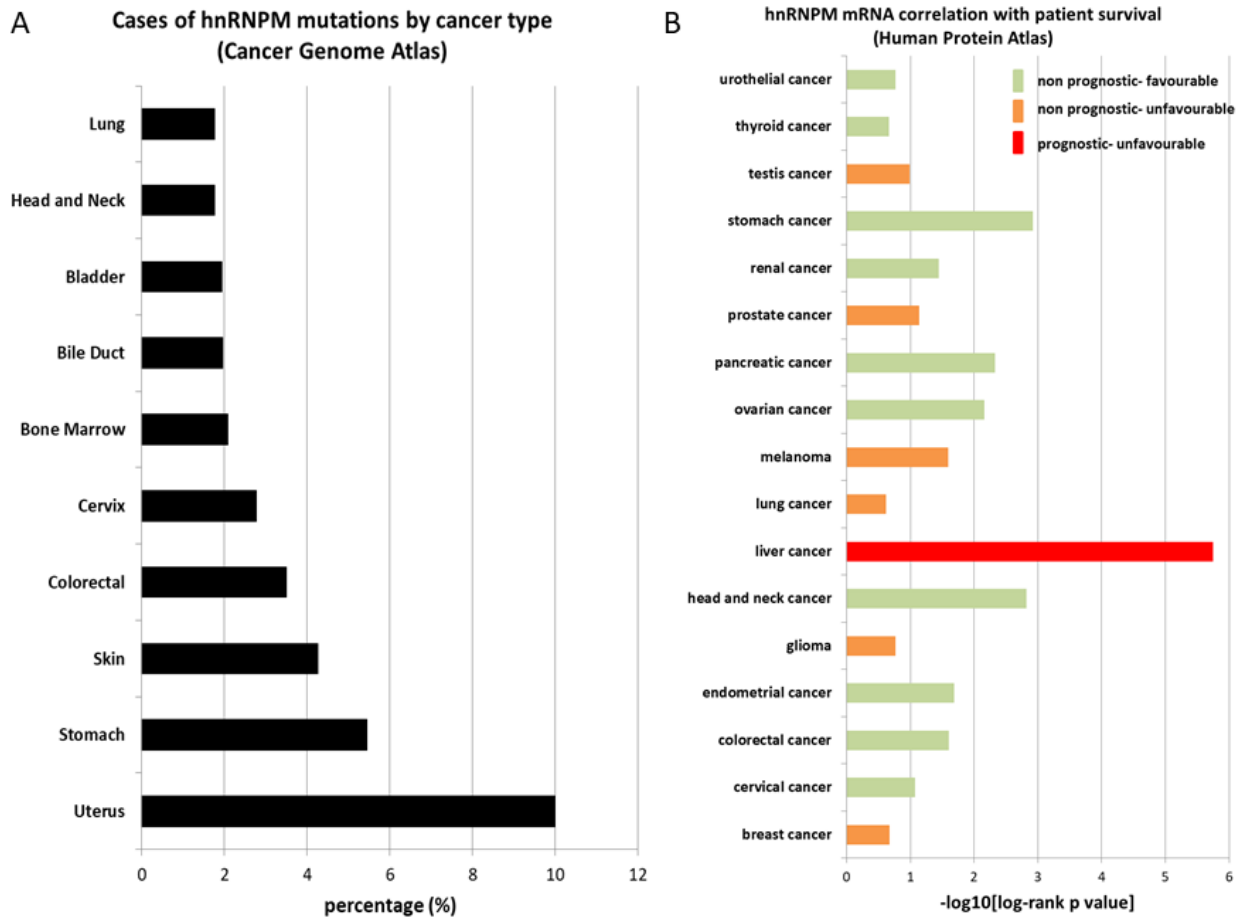

Figure 8. Association of hnRNPM with human cancers. A. The ten tumor sites which have the most cases affected by hnRNPM mutation (projects sorted by percentage of hnRNPM mutation-affected cases per total of cancer cases). B. Pathology data from the Human Protein Atlas: Prognostic value of hnRNPM mRNA expression is indicated by the log-rank p value, calculated from Kaplan Meier plots of patient survival. Log-rank p value lower than 0.001 indicates prognostic value (either favourable or unfavourable).

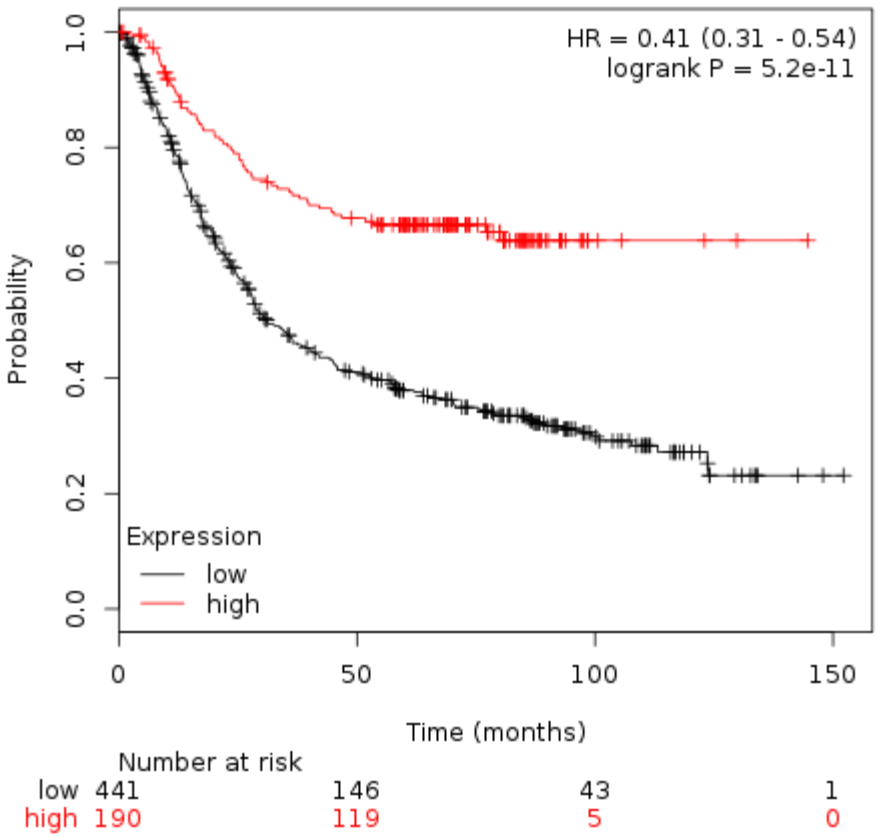

Figure 9. Kaplan-Meier plot for hnRNPM mRNA expression in gastric cancer cases. The plot was generated from publically available tissue microarray analysis data, using the online tool compiled by Gyorffy et al[199,201].

1.3.2.2 The nuclear interaction of hnRNPM with IQGAP1, a known scaffold protein

Because of hnRNPM's prominent role in spliceosomal assembly and AS regulation, and the connections between hnRNPM function and malignancy, it became a target of investigation, in an effort to elucidate mechanisms of spliceosomal assembly and AS regulation and their connection with cancer pathogenesis. In the process of this investigation, it was discovered that hnRNPM presented a direct, non-RNA-dependent protein-protein interaction with IQGAP1, a scaffold protein with known roles in regulating the assembly of numerous signaling pathways, as well cell-cell adhesion, cell proliferation and migration in the nucleus of gastric cancer cells (P. Kafasla, unpublished results). As will be discussed, IQGAP1 plays a significant role in the orchestration of signaling pathways and cellular processes, and is also implicated in cancer pathogenesis.

This novel hnRNPM-IQGAP1 interaction has important implications for the two questions outlined above, emerging from the gaps in the current research landscape: Firstly, could IQGAP1 be the missing link necessary for the connection between signaling pathways and AS regulation, and secondly, does IQGAP1-mediated regulation of hnRNPM-dependent splicing have an effect on cancer pathogenesis? The current knowledge on IQGAP1 function, which will be outlined in the following section, painted it as an extremely promising candidate for being the "hidden hand" behind the organization of AS and spliceosomal assembly. Therefore, the investigation of the novel IQGAP1-containing RBP complex and the effect of the IQGAP1-hnRNPM interaction on cancer phenotype and AS regulation formed the basis of the present work, as will be discussed in the following sections.

1.4 IQGAP1

1.4.1 Function and cellular localization of IQGAP1

IQGAP1 is a 195 kDa protein which was initially identified as an effector of Rho-GTPases, bearing a RasGAP-related domain (GRD), a calponin homology domain (CHD), an IQ domain similar to that of unconventional myosins, the WW domain- a poly-proline binding domain originally identified in dystrophin and YAP protein- and a calmodulin binding domain.[202,203]. These are domains common to all three members of the IQGAP family, which also include IQGAP2 and IQGAP3. The three members of the IQGAP family have sequence homology, but exhibit differences in their tissue and subcellular distribution and functions in health and disease (reviewed in[204–208]).

IQGAP1 also bears binding domains for extracellular signal-regulated kinase 2 (ERK2), myosin essential light chain, S100B, β-catenin[209], CLIP-170[210] and adenomatous polyposis coli (APC)[211]. The CHD domain mediates IQGAP1 interaction with F-actin[210,212], and the GRD enables IQGAP1 to act as an effector to cdc42 and Rac, members of the Rho-family GTPases (Figure 10) .[213] Rho-GTPases are a group of regulatory molecules which cycle through GTP-bound active and GDP-bound inactive states, and whose activity is required for cadherin-mediated cell-cell adhesion, cell polarization and migration. As a result of its role in modulating the activity of Rho-GTPases, IQGAP1 is a key regulator of cytokinesis, cell migration and polarization and cell-cell adhesion [213–216].

In addition to its role in the regulation of the cytoskeleton, IQGAP1 acts as a protein scaffold for signaling molecules (reviewed in[207,217]). Scaffold proteins are modular proteins equipped with multiple functional domains which enable them to recruit and assemble diverse proteins through tethering or allosteric regulation, bringing together the components of signaling or functional complexes.[218–220] Important signaling cascades orchestrated by interactions with IQGAP1 include the mitogen-activated protein kinase (MAPK) pathway. The MAPK pathway signaling cascade is activated by stimulation of growth factor receptors, leading to the sequential phosphorylation of Raf, MAPK-ERK (MEK) and ERK kinase. IQGAP1 regulates MAPK signaling by acting as a scaffold for several MAPK components, including K-Ras[221] , B-Raf [222,223], MEK[224], and ERK[224,225]. The exact mechanism of ERK1/ERK2 binding to IQGAP1 has recently been put into question, with evidence suggesting that the WW domain is not the one responsible for the ERK/IQGAP1 interaction.[226] A recent report implicates IQGAP1 in the C6 ceramide mediated activation of ERK1/2, while proposing that C6 ceramide increases IQGAP1 protein levels through preventing its cleavage by inducing acetylation of potential cleavage sites.[227] Through its interactions with β-catenin and DVL protein, IQGAP1 is also a direct regulator of the Wnt pathway.[228,229], which it also influences through cadherin mediated signaling.[207]

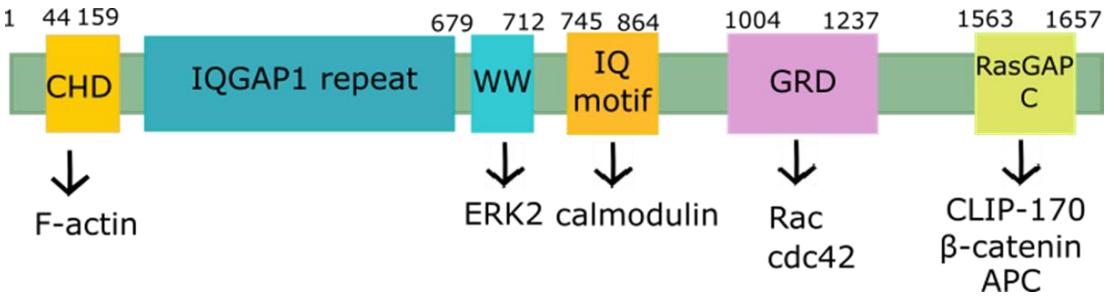

Figure 10. Structure and functional domains of IQGAP1, including the main partners of each binding domain.

As is expected, given its primary role in cytoskeletal organization, IQGAP1's predominant subcellular localization is at the plasma membrane.[207] However, IQGAP1 was shown to translocate to the nucleus during late G1-early S phase in human cell lines.[230] Nuclear localization of IQGAPs had previously been observed, with IQGAP2 being detected in the nucleus of gastric mucosal cells[205] and XTC fibroblast cell and embryonic cells.[204] IQGAP1 was reported to participate in the formation of the nucleolar envelope in murine oocytes.[231] An N-terminal fragment of IQGAP1 exhibited exclusive nuclear localization in L-cells[232], while the N-terminus of Rng2, the S.pombe homologue of IQGAP, accumulates in the nucleus upon over-expression.[233] In addition, exposure of human podocytes to puromycin aminonucleoside induced ERK-dependent nuclear translocation of IQGAP1, leading to the interaction of IQGAP1 with chromatin and Histone H3.[234]

Increasing evidence has begun to point towards a nuclear role for IQGAP1: Many of IQGAP1's known binding partners including actin, Rac1, and APC have described nuclear functions[207], and many others are nucleo-cytoplasmic shuttling proteins, including ERK, β-catenin and DVL. In fact, IQGAP1's nuclear translocation has been shown to modulate β-catenin and DVL nuclear localization.[228,229] In addition, IQGAP1 binds with and modulates the function of transcription factors, including estrogen receptor a (ERa)[235] , nuclear factor-erythroid-related factor 2 (Nrf2)[236], and nuclear factor of activated T cells 1 (NFAT1).[237] In the latter case, IQGAP1 forms part of large cytoplasmic RNA-protein complex containing NFAT1, IQGAP2, calmodulin and three NFAT kinases, which sequesters NFAT1 in its inactive state and regulates its nuclear import. Despite these indications for a nuclear function of IQGAP1, its role in the nucleus remains largely unexplored.

Interestingly for our hypothesis of IQGAP1's involvement in mRNA splicing, there is also evidence of IQGAP1's interaction with components of the mRNA decay machinery. Specifically, IQGAP1, along with its partners Rac1 and cdc 42 is a component of Staufen-containing granules.[238] Staufen is an mRNA binding protein that participates in mRNA localization and decay, through the Staufen-mediated decay (SMD) pathway, in co-operation with Upf-1, a component of the nonsense-mediated decay (NMD) pathway.[239,240]The NMD complex is responsible for the elimination of mRNA transcripts that contain premature stop codons prior to translation.[241,242] IQGAP1 also directly binds to another component necessary for NMD complex formation, SMG-9, in its dephosphorylated state. IQGAP1 binding to SMG-9 appears to inhibit NMD complex formation.[243] These studies are supportive of a role for IQGAP1 in mRNA localization and metabolism.

1.4.2 IQGAP1 and cancer

Unsurprisingly, given its role in regulating processes whose dysregulation is a significant milestone of malignant transformation of cancer cells, and its activity as a scaffold for known oncoproteins including B-raf and K-Ras, the role of IQGAP1 in cancer pathogenesis has been extensively studied (reviewed by[206,244,245]). So far, evidence suggests that IQGAP1, in contrast to its homologues IQGAP2 and IQGAP3, is an oncogene.[246] Evidence to support this proposition include the fact that increased expression of IQGAP1 has been observed in genetic studies of several human neoplasms, including lung[247], colorectal[248] and oligodendroglioma[247], as well as mouse models recapitulating human cancer phenotypes, such as prostate cancer.[249]

The GDC data on IQGAP1 include 266 cases affected by 312 IQGAP1 mutations across 27 projects.[195] The most highly affected disease type appears to be uterine corpus endometrial adenocarcinoma, with 60 out of 530 archived cases bearing an IQGAP1 mutation (11.32%), while the colon, skin, cervix, bladder, stomach, head and neck and lymph nodes are also in the top-10 primary tumor sites most highly affected by IQGAP1 mutations (Figure 11A). Data from the pathology database of the Human Protein Atlas indicate that most malignant cells showed moderate to strong staining for IQGAP1 protein expression, across all cancer types, while low-grade tumors were negative. Additionally, IQGAP1 mRNA expression is present in all cancer types, with the lowest levels in liver cancers and the highest in stomach, breast and thyroid. Regarding its effect on prognosis, IQGAP1 mRNA expression appears as an unfavourable prognostic marker for patient survival in pancreatic cancer (log-rank p value = 5.67x10-5, 176 cases), but is a weaker unfavourable indicator in urothelial, testis, or cervical cancer, and even a weakly favourable indicator (non-prognostic) in breast, endometrial, or stomach cancer (Figure 11B). This data is much less conclusive than individual projects examining IQGAP1 expression in cancer, such as those mentioned above, regarding the levels of IQGAP1 mRNA or protein over-expression and its overall effect on human malignancy and disease prognosis. However, the fact remains that IQGAP1 mutations or changes in gene expression have been recorded in large scale genetic studies of human neoplasms.

The mechanistic role of IQGAP1 in the process of tumorigenesis is also supported by data from cultured human cancer cell lines: IQGAP1 amplification of mRNA and protein levels has been observed in the gastric cancer cell lines HSC39 and HSC40A.[250] In addition, silencing of IQGAP1 using si-RNA and shRNA techniques has been shown to inhibit invasion, cell mobility and migration of HO-8910PM cells *in vitro* [251] and the migration/invasion of MCF-7 cells[246] and U87MG human glioblastoma cells[252] *in vitro*. At the same time, over-expression of IQGAP1 in MC-7 cells enhances migration and proliferation both *in vitro* and *in vivo*.[246] IQGAP1 post-translational modifications through signaling pathways have an important role in its function, and appear to influence its role in cancer progression: it has been suggested that IQGAP1 SUMOylation promotes colorectal cancer progression *in vitro* and *in vivo*.[253]

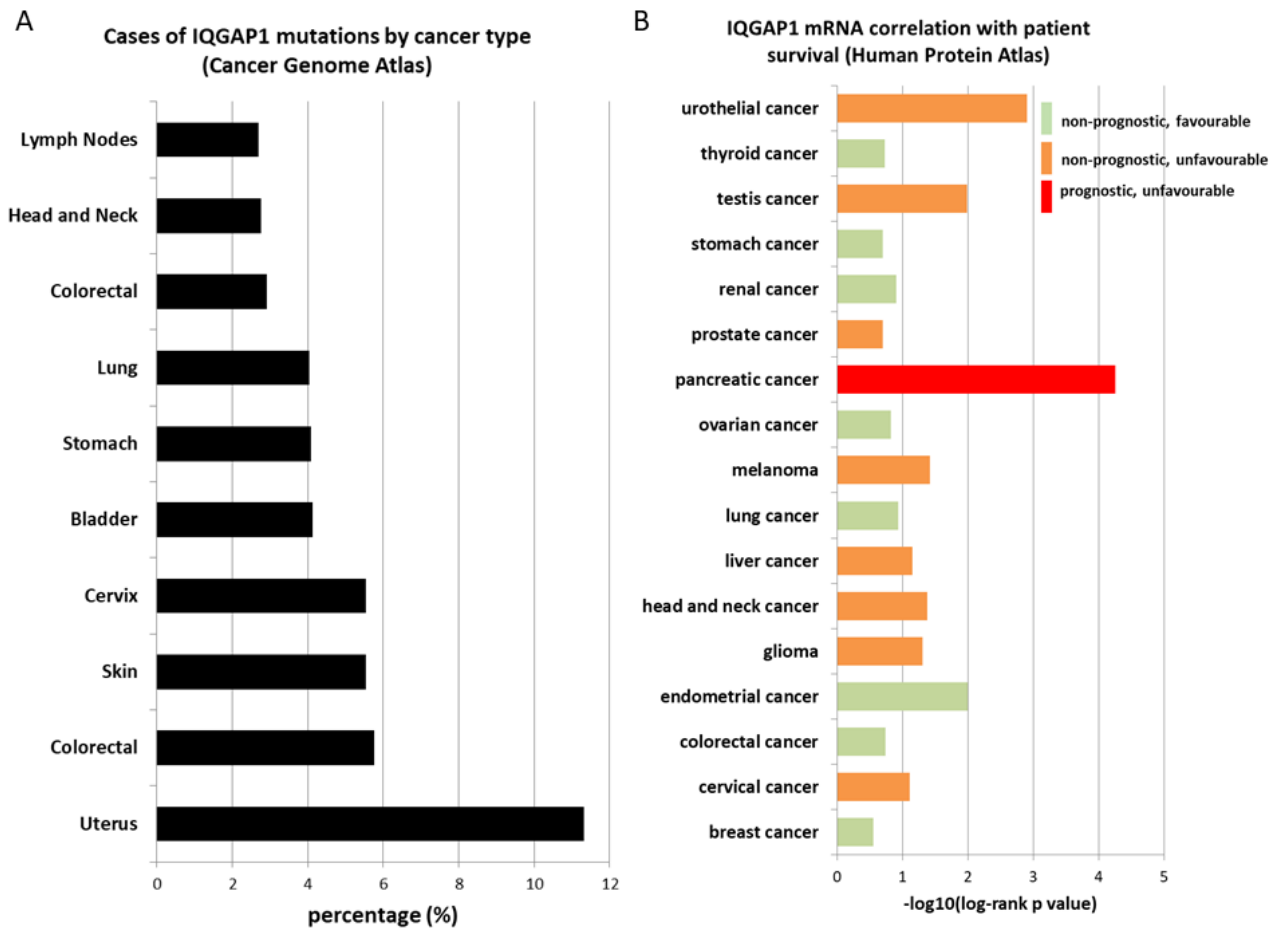

Figure 11. Association of IQGAP1 with human cancers. A.The ten tumor sites which have the most cases affected by IQGAP1 mutation (projects sorted by percentage of IQGAP1 mutation-affected cases per total of cancer cases). B. Pathology data from the Human Protein Atlas: Prognostic value of IQGAP1 mRNA expression is indicated by the log-rank p value, calculated from Kaplan Meier plots of patient survival. Log-rank p value lower than 0.001 indicates prognostic value (either favourable or unfavourable).

1.4.3 Implications of IQGAP1's interaction with hnRNPM

Both hnRNPM and IQGAP1 are master regulators of processes whose dysregulation may well lead to or promote cancer pathogenesis. The novel IQGAP1-hnRNPM interaction, therefore, presented great interest as an investigative starting point, emerging as a potential answer to both the questions outlined in the previous sections .Firstly, the interaction of hnRNPM, a known orchestrator of AS, with a scaffold protein, known to coordinate numerous signaling processes has important implications for the elucidation of AS regulation through signaling pathways. Because of its ability to bring together proteins and participate in signaling pathways, including the MAPK and PI3k-Akt signaling cascades, IQGAP1 could be the missing link between extracellular or intracellular stimuli and the differential assembly or post-translational modifications of spliceosomal components which are known to lead to AS events (see 1.2.2). Secondly, this potential role of the IQGAP1-hnRNPM interaction in the regulation of AS could help shed light on the as-yet unexplored nuclear role of IQGAP1, and expand the mechanisms of the involvement of both proteins in cancer pathogenesis, with a view towards a potential target for therapeutic intervention. Initial investigation of this novel protein-protein interaction unearthed RNA-dependent interactions of IQGAP1 with other spliceosomal components (GR Manikas, P Kafasla, unpublished results), leading to the hypothesis that a novel RNP complex is formed, with the hnRNPM-IQGAP1 interaction at its heart.

In order to investigate the role of this novel complex in cancer pathogenesis, as well as splicing regulation, human gastric adenocarcinoma cultured cell lines was selected as a model system, prior to the initiation of this thesis. There are three main reasons informing the choice of gastric cancer as a focus of study:

1. The prognostic correlation between hnRNPM RNA expression and favourable prognosis in gastric cancer patients, which was not observed in breast, lung or ovarian cancers, cancers that have also been shown to be affected by alternative splicing changes in general (Sveen et al, Oncogene, 2016) and hnRNPM expression in particular. This is also supported by the fact that, as previously mentioned, 5.45% gastric adenocarcinoma cases are affected by hnRNPM mutations.
2. The knowledge that IQGAP1 knockout mouse model exhibits a phenotype of gastric hyperplasia[254], indicating a significant role of IQGAP1 mediated functions in specific gastric tissue development. At the same time, data from gastric cancer cell lines suggests an inverse correlation between E-cadherin adhesion and IQGAP1 localization.[255] IQGAP1 appears to be differentially expressed in different sub-types of gastric epithelial cells in rabbit-derived primary culture, predominantly targeted to the apical poles of chief and muckous neck cells through protein kinase C-dependent signaling.[205] This data suggests IQGAP1 may have a physiological role for gastric tissue development and function.

Chapter 2. Aims of the project

Alternative splicing is a post-transcriptional regulatory process of gene expression, which affects nearly all genes.[14,15,21] Recently, it has emerged as an underlying mechanism driving tumour growth, invasion and metastasis, and designated a "hallmark of cancer".[156] Elucidation of the mechanisms underlying AS regulation in cancer would be a critical step towards understanding the process of tumor development and progression, and developing novel therapeutics.[157]

Regulation of AS is orchestrated by the splicing regulating proteins, which are organized and modified into distinct functional ribonucleoprotein complexes (RNPs). Assembly of the spliceosomal complexes responds to changes in environmental or cellular conditions through extra- or intra-cellular signals.[106,112,114] However, there are many unanswered questions concerning the mechanism through which signaling pathways regulate AS.

The aim of the present master's thesis was to investigate a novel finding of the host lab, the interaction of the scaffold protein Iqgap1 with the splicing regulator hnRNPM in the nucleus of human gastric cancer cell lines (Kafasla P, preliminary data). This finding presents great interest, given that Iqgap1 has an extensive, well documented role in mediating signal transduction in the cytoplasm, but very little is known about its role in the nucleus or in AS regulation, suggesting this finding may have important implications.
The lab's working hypothesis was that Iqgap1 plays a significant role in post-transcriptional regulation of gene expression, by facilitating the assembly of functional RNP complexes and mediating post-translational modifications of splicing regulators, including hnRNPM. The aim of this project was to contribute to the investigation of the function of this novel IQGAP1-containing RNP and its role in AS regulation and gastric cancer development and progression.

To this end, protein co-immunoprecipitation and mass-spectrometry proteomic techniques were employed in normal gastric epithelial and gastric cancer cell lines in order to identify nuclear proteins interacting with IQGAP1 in both wild type and disease conditions, and study their molecular and biological functions and post-translational modifications, seeking a confirmation of the preliminary results suggesting IQGAP1's interaction with various spliceosomal components. Additionally, gene-editing techniques were employed in order to generate IQGAP1 knock-out cell lines, as a first step in investigating the combined phenotype of IQGAP1-hnRNPM depletion in normal and gastric cancer cells. Finally, the specific role of the IQGAP1-hnRNPM interaction is splicing regulation was investigated using an hnRNPM-dependent splicing assay on IQGAP1-depleted cells. The questions we aimed to answer with this work were:

1. Whether IQGAP1 has a mechanistic function in the regulation of alternative splicing, and
2. Whether this function influences characteristics related to the development and progression of the gastric cancer phenotype, including alternative splicing.

Chapter 3. Materials and methods

3.1. Cell culture

Two established cell lines were used: an immortalized gastric epithelial cell line, HFE-145 originally generated from endoscopically biopsied gastric epithelial cells via transfection with SV40 T-Ag and human telomerase (hTRT-145) [256], and a gastric cancer cell line originating from the lymph nodes of a poorly differentiated metastatic gastric adenocarcinoma, NUGC-4.[257]All cell lines were appropriately cultured at 37ºC, in a humidified atmosphere containing 5% CO_2. HFE-145[256] cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Merck) with 10% (v/v) fetal bovine serum (FBS, BioSera), 100 U/mL penicillin/streptomycin (Gibco) and 2 mM L-glutamine (Gibco) and buffered in 5 mM NaOH, 0.2% sodium bicarbonate. NUGC-4[258,259] cells were cultured in Roswell Park Memorial Institute Medium (RPMI, Gibco) supplemented with 10% v/v FBS, 100 U/mL penicillin/streptomycin (Gibco) and 2 mM L-glutamine (Gibco).

3.2. Preparation of nuclear and cytoplasmic extracts

The protocol for sub-cellular fractionation was adapted from Dreyfuss et al, PNAS (1984)[25]. For each experiment, approximately $1.0x10^7$- $1.0x10^8$ cells were appropriately harvested. Cells were immediately used or re-suspended in 50% glycerol/phosphate buffered saline (PBS) solution and stored at -80ºC. Washes of the isolated cells were performed with centrifugation at 100xg/3 min at 4ºC with cold PBS. The cell pellet was re-suspended in 3 to 5 volumes of hypotonic Buffer A (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 2.5 mM MgCl2) supplemented with 0.5% Triton X-100, protease and phosphatase inhibitors (NaF 1mM, Na_3VO_4 1 mM) and incubated on ice for 5-10 min. Cell membranes were sheared by passing the suspension through a 26-gauge syringe 4-6 times. The nuclei were isolated with centrifugation at 3000xg for 10 min at 4ºC, and the supernatant, which is the cytoplasmic extract, was aliquoted and retained at -80ºC. To avoid contamination with cytoplasmic components, the nuclear pellet was washed by resuspension in the same buffer and centrifugation as before. Following this step, the nuclei were re-suspended in 2 volumes of Buffer A and sonicated twice for 5s (0.2A), with a 15s rest on ice in-between sonications. The sample was centrifuged at 4000xg for 10 min at 4ºC. The upper phase, which is the nuclear extract, was collected, aliquoted and stored at -80ºC, while the nuclear pellet was re-suspended in 2 volumes of 8 M Urea and stored at -20ºC. Protein concentration of the isolated fractions was assessed using the Bradford assay.[260]

3.3. Immunoprecipitation of proteins

Co-immunoprecipitation of proteins was performed on nuclear extracts, using agarose beads obtained by SantaCruz Biotechnology (Protein A/G Plus Agarose Beads, sc-2003). Beads were prepared for use as follows: 20 ul of bead slurry per immunoprecipitation reaction was washed with NET-2 buffer (10 mM Tris, pH 7.5, 150 mM NaCl, 0.05% NP-40) and blocked in 3% BSA in NET-2 (0.5 mL) for 30 min at 4ºC on a rotating wheel. Blocking buffer was discarded after centrifugation at 100xg/4ºC for 3 min. For antibody binding to the beads, 4-8 ug of antibody per reaction was added, to a final volume of 500-600 uL in NET-2 buffer. Antibody binding was performed by overnight incubation at 4ºC on a rotating wheel. Following the binding of the antibody, beads were washed at least 3 times by resuspension in NET-2 buffer (after each washing step, suspended beads were recovered by centrifugation, as described). The appropriate amount of antibody-bound beads was allocated to each immunoprecipitation reaction, suspended in NET-2 buffer. The

sample binding to the beads was performed as follows: Prior to usage, the nuclear extract was centrifuged at 10.000 rpm for 1 min at 4ºC, and any precipitate was discarded. The appropriate volume of nuclear extract (500-1000 ug of protein) was added to each reaction to a final volume of 600-800 uL in NET-2 buffer. The sample was incubated with the beads for 2h at 4ºC on a rotating wheel. After sample binding, beads were recovered as before, and an aliquot of the supernatant was kept to check flow-through of the sample. Sample-bound beads were washed 3 times with NET-2 buffer, and twice with NET-2 buffer supplemented with 0.1% Triton X-100 and a final concentration of 0.1% NP-40. After the final washing step, care was taken to completely remove the supernatant. Co-immunoprecipitated proteins were eluted from the beads by adding 15-20 uL of 2x Laemmli sample buffer (0.1 M Tris, 0.012% bromophenol blue, 4% SDS, 0.95 M β-mercapthoethanol, 12% glycerol) and boiled at 95ºC for 5 min. Following centrifugation at 10.000 x g for 2 min, the supernatant was retained and stored at -20ºC or immediately used. The quantities of total protein for nuclear extracts used for immunoprecipitation reactions ranged from 300-1000 μg. Antibodies used for immunoprecipitation reactions are listed in supplementary table 1 (Addendum).

3.4. Western blotting

Protein extracts, were reduced by boiling at 95ºC for 5 min in 1x Laemmli sample buffer. For effective visualization of IQGAP1 in nuclear extracts, 30 ug of protein sample was required. Samples were run in 8-10% SDS-PAGE at 80-120V using SDS running buffer (1% SDS, 192 mM glycine, 25 mM Tris base) or in pre-cast gels (4-12%, Invitrogen) in Invitrogen X-cell Sure-Lock at 130 V using MOPS buffer (20 mM MOPS, 5 mM sodium acetate, 1 mM Na_2EDTA). Gels were transferred to PVDF membrane (GE Healthcare, 10600023) for 2h, at 400 mA, using 1x transfer buffer (1% SDS, 192 mM glycine, 25mM Tris base, 15% methanol). Membranes were probed with appropriate primary and secondary antibodies. The list of primary and secondary antibodies used for immunoblotting in this study is included in supplementary table 1 (Addendum).

3.5. Mass spectrometry and Proteomics analysis

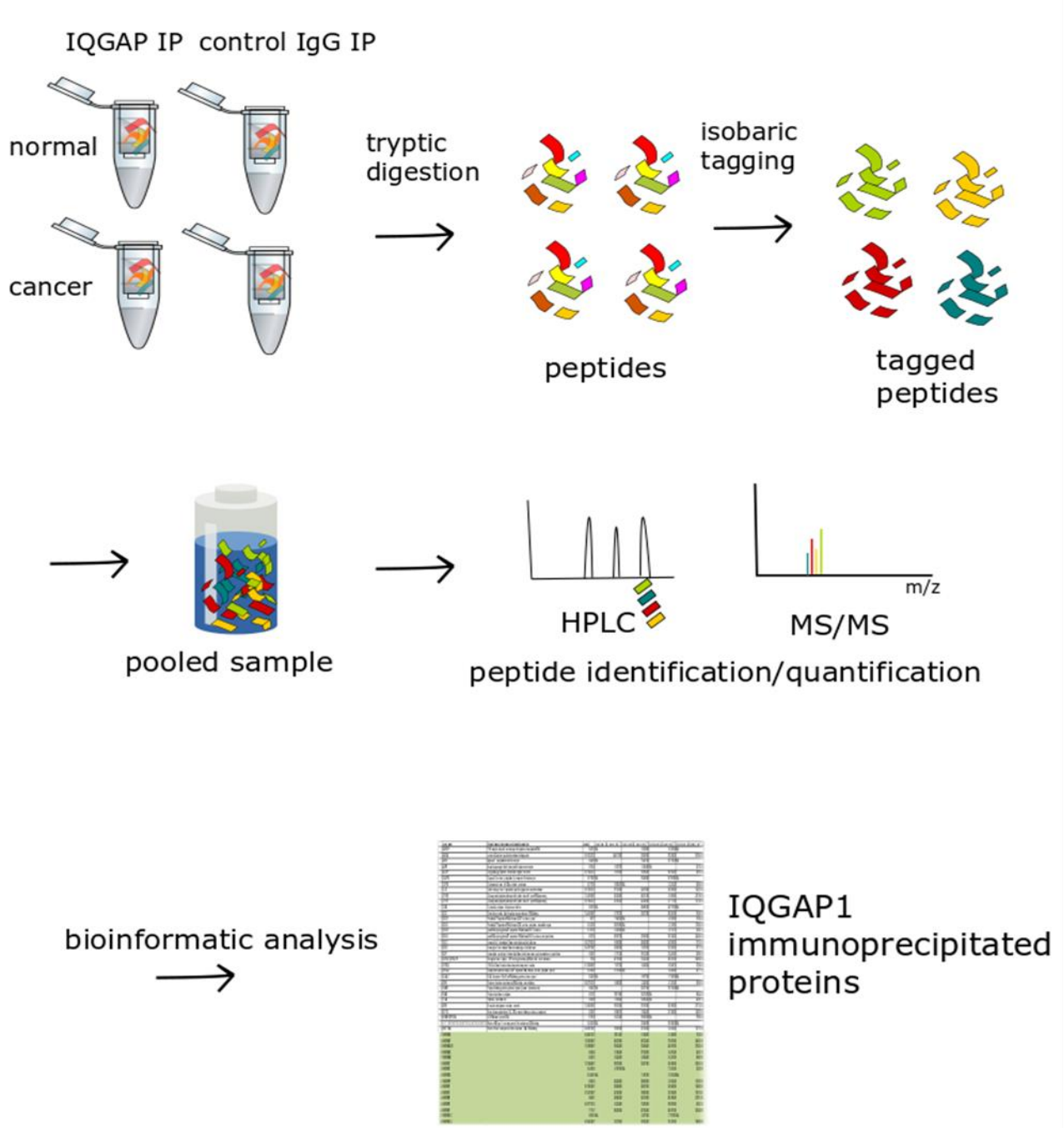

Figure 12. Experimental procedure for proteomics analysis of IQGAP1 IPs

IP samples were processed in collaboration with the Core Proteomics Facility at EMBL Heidelberg. Proteomics analysis[9] was performed with the following protocol (Figure 12). The samples, dissolved in 2x Laemmli sample buffer, as described, and underwent filter-assisted sample preparation (FASP) to produce peptides with proteolytic digestion.[261] These were then tagged using 4 different multiplex TMT isobaric tags (ThermoFisher Scientific, TMTsixplex™ Isobaric Label Reagent Set): one isotopically unique tag for each IP condition, namely IQGAP1 IP cancer (NUGC-4), normal (HFE), and their respective IgG controls. TMT-tagged samples were appropriately pooled and analyzed using HPLC-MS/MS. Three technical replicates for each IP condition were processed, resulting in 3 pooled replicates.

TMT-tags are isobaric compounds, ie compounds of the same mass and structure, but which vary in the distribution of heavy isotopes in their structure.[262] These isobaric tags are composed of three structural elements: a reporter, a balance and a reactive region. The reactive region is the same in all tags, and consists of a primary amine-reactive N-hydroxysuccinimide (-NHS) group. This region is responsible for attaching the tag to peptides produced by tryptic digestion through irreversible, covalent binding of their NHS-ester activated groups to peptide primary amine groups ($-NH_2$ groups). Each digested experimental sample reacts with a different isotopic variant of the tag from a set, and the samples are mixed in equal ratios and analyzed simultaneously in one MS run. Because the tags are isobaric and have identical chemical properties, tagged identical peptides from each sample have identical HPLC retention times and appear as a single precursor ion composite peak in the first MS scan. The reporter and balance regions differ in isotopic distribution between each tag in a multiplex set, so that each reporter region produces peaks of characteristic m/z ratio after high energy collision induced dissociation (MS2). This allows the simultaneous identification and comparative quantification of identical peptides (and by consequence, proteins) in each of the multiplexed samples. The simultaneous processing of multiplexed samples results in reduced processing variability and increased specificity in peptide quantification.

Samples were processed using the ISOBARQuant[263], an R-package platform for the analysis of isobarically labelled quantitative proteomics data. Only proteins that were quantified with two unique peptide matches were filtered. After batch-cleaning and normalization of raw signal intensities, a fold-change was calculated. Statistical analysis of results was performed using the LIMMA[264] R-package, making comparisons between each IQGAP1 IP sample (cancer, normal) and their respective IgG controls (cancer, normal), as well as between IQGAP1 IPs (cancer vs normal). A protein was considered significant if it had a p-value < 5% (Benjamini-Hochberg FDR adjustment), and a fold-change of at least 50% between compared conditions. Identified proteins were classified into three categories: Hits (FDR threshold= 0.05, fold change=2), candidates (FDR threshold = 0.25, fold change = 1.5), and no hits.

3.6. Analysis of phosphorylation in IQGAP1-immunoprecipitated proteins

Preparation of nuclear extracts and immunoprecipitation using an anti-IQGAP1 primary antibody was performed as above. The samples (IQGAP1 IPs in HFE-145, NUGC-4 and the respective normal rabbit IgG controls) were subsequently split in half and processed with the following sample preparation protocols:

1. Filter-Assisted Sample Preparation, performed as previously described[261]. Briefly, samples were loaded to filter-tubes and denatured with urea (7 M, 300 µl). After centrifugation (14.000 xg, 20-45 min) thiols were carboxyamidomethylated with addition of iodoacetamide (IAA) (10 mg/mL, 100 µl) and incubated in the dark for 15'. Excess reagent was removed by centrifugation (as before). Urea denaturation was repeated three more times. Peptides were washed with addition of 0.2% ammonium bicarbonate solution (100 µl) , followed by centrifugation. The washing step was repeated three times. Filters were moved to a fresh tube, and 1 µg of trypsin in ammonium bicarbonate 0.2% was added per sample. Samples were incubated O/N at 37ºC, 300 rpm, in the dark. The following day, samples were centrifuged at 14.000xg for 10 min. Double distilled water for mass spectrometry was added to the filters and samples were incubated on a shaker for 1 h (at 37ºC, 300rpm). Peptides were eluted by centrifugation at 14.000xg for 30 min, and stored at -30ºC. Prior to usage, samples were dried using a SpeedVac vacuum concentrator for 1-4 hours (depending on sample concentration). Dehydrated peptides can be stored or dissolved in buffer A (2% acetonitrile, 0.1% formic acid), for use. Dissolved peptides were sonicated for 2-3 minutes in a water bath, and subjected to a quick centrifugation. Sample concentration was determined by measuring absorbance at 280 nm on a Nano-drop.

2. Single-Pot-Solid Phase Enhanced Sample Preparation (SP3), adapted from a previously described protocol.[265]A mix of equal volumes of GE Healthcare Sera-Mag SpeedBeads™ Carboxyl Magnetic Beads, hydrophobic (65152105050250) and hydrophilic (45152105050250) were used. Beads were prepared by combining 50 µl of each bead solution, placing on a magnetic rack and letting them settle for 2 min. Supernatant was discarded, and beads rinsed by resuspension in 200 µl of 50 mM acetic acid, for three steps. Prior to use, beads were stored in 500 µl of acetic acid solution at 4ºC. IP samples (350 µg initial nuclear extract per sample) were eluted from beads in eluted in lysis buffer (100 mM Tris-HCl pH 7.6, 4% SDS, 0.05 M DTT). Sample were diluted (with lysis buffer) to a final volume of ~90 µl, and 7.5 µl of iodoacetamide (200 mM) was added to a final concentration of 15 mM, for the thiol alcylation reaction. Samples were incubated in the dark for 30 min. The reaction was quenched by addition of 10 µL DTT 200 mM. Bead solution was added to the sample (2 µL of bead solution per sample). 100% Acetonitrile/1% formic acid solution was added to obtain a final acetonitrile concentration of 50% (~100 µl). Samples were mixed gently with pipetting, and incubated for 8 min in a rotating rack. Tubes were placed in a magnetic rack and incubated for 2 min, and the supernatant was discarded. 200 µl of 70% ethanol was added and the samples incubated for 30s in the magnetic rack, after which the supernatant was discarded. The washing step was repeated, and then 180 µl of acetonitrile 100% was added, and samples incubated for ~ 1 min in magnetic rack. Supernatant was discarded, and the protein-bound beads air-dried for ~2 min. Beads were reconstituted in 0.05 µg/µL LysC/Trypsin solution in 100 mM ammonium bicarbonate (0.5 µg per sample), and mixed gently by pipetting. Samples were sonicated for 5 min in a waterbath and incubated at 37ºC O/N with mild shaking, for on-bead tryptic digestion. The following day, samples were subjected to a quick centrifugation, to recover solvent, and resuspended in an added 50 µL of H2O, followed by a 2 min sonication to dissolve digested peptides. Samples were placed on the magnetic rack for ~ 2min. At this point, the magnetic beads were expelled from the aqueous supernatant, which was removed and transferred to low-protein binding PCR tubes. Samples were dried in a Speedvac vacuum concentrator and reconstituted for HPLC- MS/MS injection in Buffer A.

Samples prepared with FASP and SP3 protocols were processed in sequential runs. The purified peptides were analyzed by HPLC (Ultimate 3000, ThermoFisher Scientific) coupled to an LTQ- Orbitrap XL Mass spectrometer (ThermoFisher Scientific, USA) equipped with a nanospray source. Ten µl of the peptide mixtures were pre-concentrated at a flow-rate of 3µl/min for 10 min using a C18 trap column (Acclaim PepMap100, ThermoFisher Scientific) and then loaded onto a 50 cm C18 column (75 µm ID, particle size 2 µm, 100Å, Acclaim PepMap RSLC, Thermo Scientific). The binary pumps of the HPLC (RSLCnano, Thermo Scientific) contained solution A (2% (v/v) acetonitrile in 0.1% (v/v) formic acid) and solution B (80% acetonitrile in 0.1% formic acid). The peptides were separated using a linear gradient of 4- 40% B in 110 min at a flow rate of 300 nl/min. The column was placed in an oven operating at 35°C. Full scan MS spectra were acquired in the orbitrap (m/z 300–1600) in profile mode and data-dependent acquisition with neutral loss activated, with the resolution set to 60,000 at m/z 400 and automatic gain control target at 10^6. The six most intense ions were sequentially isolated and subjected to collision induced dissociation (CID) and detection in the linear ion-trap to acquire the MS/MS spectra. Dynamic exclusion was set to 60 sec. Ions with single charge states were excluded. Lockmass of m/z 445,120025 was used for internal calibration. The software Xcalibur (Thermo Scientific, USA) was used to control the system and acquire the raw files.

3.7. Generation of cell lines with stable IQGAP1 knockdown by CRISPR-Cas9

The CRISPR/Cas9 strategy was used to generate IQGAP1 knockout cells. The method is implemented in mammalian cells by co-expressing the bacterial Cas9 enzyme with the synthetic guide RNA[266] (Figure 13A). The synthetic guide RNA has a sequence complementary to the target gene sequence and a scaffold se-

quence, necessary for binding to Cas9. The target sequence is usually ~ 20 bp long and must be flanked by the protospacer adjacent motif (PAM) which is recognized by the Cas9 enzyme. After the sgRNA and Cas9 are both expressed, the enzyme forms a ribonucleoprotein complex with the guide RNA (sgRNA:Cas9 complex), which is able to bind genomic DNA, while the sgRNA remains free to interact with the target sequence. Once the sgRNA guides the complex to the target sequence, it begins to anneal to the target in a 3'-5' direction. After annealing, the presence of the PAM allows Cas9 to cleave opposite strands of the target DNA, introducing a double-strand break (DSB) within the target DNA (3-4 nucleotides upstream of the PAM sequence). This cleavage will then be repaired by one of the two following pathways:

1. The Non-Homologous End Joining (NHEJ) repair pathway
2. The Homology Directed Repair (HDR) pathway

Repair by the NHEJ pathway is the most common and efficient repair mechanism, but it frequently introduces insertions or deletions in the repaired sequence, which result in amino-acid deletions, insertions, premature stop codons or frameshift mutations. Repair of a Cas9-induced DSB via the NHEJ pathway should ideally lead to a loss of function mutation, "knocking-out" the target gene. In order to reduce the method's off-target effects and increase specificity, the bacterial Cas9 D10 nickase is used (Figure 13B)[266,267]. Cas9 D10 nickase is an enzyme which introduces single-stranded nicks instead of DSBs. As a result, two nickases targeting opposite DNA strands are required to generate a DSB within the target. Since the probability of two off-target nick being generated in close enough proximity to cause a DSB is very low, target specificity is greatly increased.

Here, the IQGAP1 gene was targeted using a pair of synthetic guide RNA (sgRNA) sequences, Assembly 1 and Assembly 2, which target Exon 1 of the IQGAP1 gene at the following sequences: Assembly 1: 5'-CACTATGGCTGTGAGTGCG-3' and Assembly 2: 5'- CAGCCCGTCAACCTCGTCTG-3' (Figure 13B). The sequences were identified using the CRISPR Design tool (Feng/ Zhang laboratory, MIT).[268] These sequences and their reverse complements were annealed and ligated into the All-In-One vector (AIO-Puro, Addgene catalogue #74630), from the laboratory of Steve Jackson, which encodes Cas9-D10A nickase linked via 2A peptide with puromycin resistant marker and dual U6 promoter guide RNA scaffolds (Figure 14), as previously described.[269]

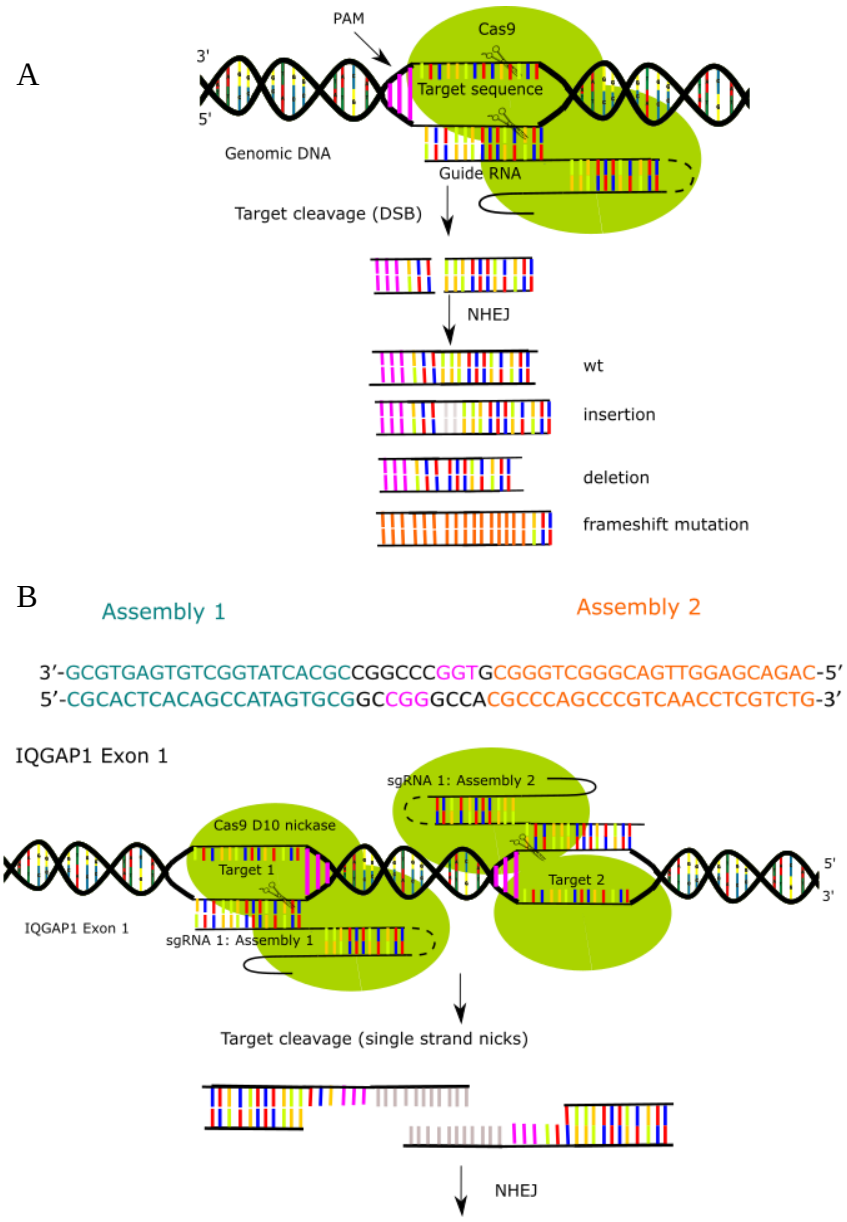

Figure 13. Principle of the employed CRISPR-Cas9 technique. A. General principle of CRISPR-Cas9 strategy.. B. Use of CRISPR-Cas9 D10 nickase to target Exon 1 of the IQGAP1 gene, using two sgRNAs (Assembly 1 and Assembly 2) targeting adjoining regions of the Exon 1 sequence.

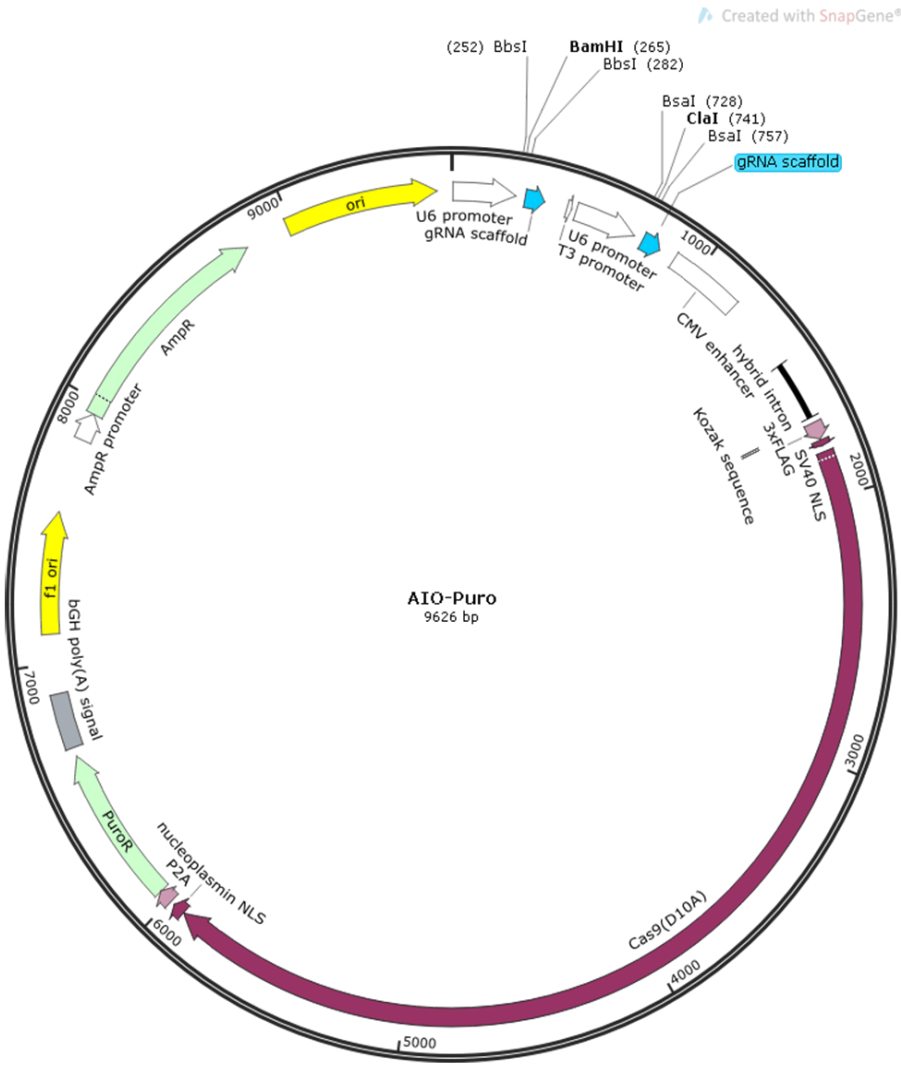

Figure 14. Sequence map of the empty vector All-in-One plasmid encoding dual U6 promoter-driven sgRNAs and Cas9-D10A nickase linked via 2A peptide with puromycin resistant marker, Addgene plasmid #74630, Steve Jackson lab[9]

AIO-Puro vector plasmids were digested with BbSI (NEB) and then dephosphorylated by Thermosensitive Alkaline Phosphatase (TSAP, Promega, catalogue number M9910). The digested vectors were recovered from 0.8% agarose gel using the NEB Monarch ™ DNA gel extraction kit and melted at 65°C for subsequent ligation reactions. The two pairs of complementary DNA oligos (Assemblies 1 and 2 including a 4-mer overhang + 20-mer of sgRNA sequence) were purchased in standard desalted format from Invitrogen. The DNA oligonucleotides were individually phosphorylated using T4 polynucleotide kinase (PNK, NEB) and then annealed by pooling together in a thermocycler machine using the following program: 95°C for 5 minutes, followed by gradual cooling from 95°C to 25°C at a ramp rate of 5°C per minute. Each DNA oligo duplex had 5' overhangs (forward: CACCG, reverse: AAAC) designed to be directly cloned into the BbSI or BsaI-digested AIO-Puro vector. The first oligonucleotide was ligated into BbsI-digested AIO-Puro vectors using T4 DNA Ligase (NEB).

The successful clones lost a unique BamHI recognition site that was confirmed by BamHI digestion alongside with a control of an empty vector. To clone the second DNA oligonucelotide duplex, the same procedure was followed, using BsaI to digest the vector. The successful clones lost a unique ClaI recognition site that was confirmed by ClaI digestion alongside a control of an empty vector. The insertion of sgRNAs was verified via sequencing. HFE-145 and NUGC-4 cells were transfected using Lipofectamine 2000, and

clones were selected 48 h later using puromycin. Individual clones were plated to single cell dilution in 24 well-plates, and IQGAP1 deletion was confirmed by PCR of genomic DNA extracted with Nucleo-Spin plasmid kit (Macherey-Nagel) using the following primers:
Forward: 5'- GCCGTCCGCGCCTCCAAG -3'
Reverse: 5'- GTCCGAGCTGCCGGCAGC-3'

Loss of IQGAP1 protein expression was confirmed by Western Blotting. HFE-145 and NUGC-4 cells transfected with AIO-Puro empty vector were selected with puromycin and used as a control during the clone screening process.

3.8. Colony formation assay

Proliferative capacity of the IQGAP1 knockout cell lines was measured using a colony formation assay for adherent cells.[270] Sub-confluent or confluent cells (controls and stable IQGAP1 knockout clones) were appropriately harvested, resuspended in fresh medium, and counted. Cells were manually diluted, as necessary, in order to acquire a concentration of at least 10 cells per microliter. The appropriate volume of cell suspension was added by pipetting, seeding each well of a 6 well plate with 100, 200 or 400 cells. Cells were cultured for 1-2 weeks at standard conditions, and checked for the formation of colonies using a bright-field microscope. Once colonies visible under low-resolution bright-field microscopy were formed in the control cells, cells were fixed as follows: medium was removed, taking care not to detach the colonies, and sufficient 100% methanol was added to cover the cells. The dish was covered and incubated at room temperature for 20 min. Methanol was subsequently removed and cells carefully rinsed with H_2O. Sufficient crystal violet stain (0.5% crystal violet in 10% ethanol) was added to cover the cells. The dish was incubated for 5 min at r.t. The cells were washed under running water until excess dye was removed, and left inverted to dry overnight. Pictures were taken and colonies counted manually. Proliferative capacity of each IQGAP1 depleted cell line was assessed by calculating the surviving fraction (SF) of each knockout based on the plating efficiency (PE) of its respective control, using the following formulas:
.

$$PE = \frac{no.of\ colonies\ formed}{no\ of\ cells\ seeded}$$

$$SF = \frac{no.of\ colonies\ formed}{no.of\ cells\ seeded\ \times PE}$$

3.9. Cell cycle analysis

Cell cycle progression in wild type and knock-out cell lines was assessed using fluorescence-activated cell sorting (FACS) analysis of propidium iodide stained cells.[271,272] Cells were cultured overnight in 6 well-plates, to a confluency of 60-70%, appropriately harvested and washed in cold PBS. At least $1x10^6$ cells were fixed in 1 mL cold ethanol (70%), which was added dropwise to the cell pellet while vortexing to minimize clumping. Cells were fixed for at least 30 min at 4°C. Cells were centrifuged at 3000 rpm for 20 min at 4°C the supernatant was carefully discarded. Cells were washed in 1 mL Phosphate-Citrate buffer (192 parts: 0.2 M Na_2HPO_4, 8 parts: 0.1M citric acid, pH=7.8) and centrifuged as before. Care was taken to avoid cell loss when discarding supernatant. To ensure that only DNA was stained, cells were simultaneously treated with Ribonuclease A. An RNase/PI solution was prepared by mixing 50 uL of 100 ug/mL RNase and 200 uL of 50 ug/mL PI for each sample. Each sample was incubated with 250 uL of PI/RNase solution (40 ug/mL PI, 20

ug/mL RNase A) for 30 min at 37ºC, and analysed by flow cytometry within one hour. Flow cytometry was performed using a BD FACS Canto™ II system, setting the G0/G1 gate of the PI-A histogram at 50 units. Cell cycle analysis was performed using FlowJo ™ software, employing a Watson pragmatic algorithm model on live singlet cells.

3.10. Mini-gene splicing assay

hnRNPM-dependent splicing reporter mini-gene assays were performed in HFE-145 and NUGC-4 cells as well as their respective IQGAP knockout derived cell lines, as previously described.[273] Briefly, cells were seeded in a 24- well plate 24 h before transfection using Turbofect transfection reagent (ThermoFisher Scientific). Cells were transfected with a three exon-two intron minigene construct based on DUP51EK[49,274] which was created in and kindly provided by D.L. Black's lab.[273] Exon 2 of the wild type reporter is included at 90%.[49] The construct used in this study was designed to detect alterations in hnRNPM-dependent alternative splicing, and the second exon contains an hnRNP M consensus binding motif, UGGUGGUG, with hnRNPM binding leading to exon-silencing.[273] An equivalent minigene was also used, carrying a mutation in the hnRNPM site (pDUP51-ΔMsite).[273] 1ug of the pDUP51 and pDUP51-ΔMsite minigene construct were transfected in cells, with or without plasmids expressing FLAG-IQGAP or EGFP-IQGAP1. Cells were collected 40 h after transfection, and RNA was isolated using TRI Reagent ® (Sigma-Aldrich) for subsequent RT–PCR analysis.

3.11. RNA isolation/Reverse Transcription/PCR

RNA isolation was performed using the TRI Reagent ® RNA/DNA/Protein isolation reagent (Sigma Aldrich, TR118) according to the manufacturer's protocol. The precipitated RNA was dissolved in ddH_2O and quantified using a Nano-drop spectrophotometer at 260 nm. Quality of isolated RNA was checked by running samples in a 1% agarose gel stained with Serva DNA Stain G or Ethidium Bromide.

Reverse transcription (RT) was performed as follows: The RT reaction was set up using 400 ng of isolated RNA per reaction, and adding 10 nmol dNTPs, 0.5 pmol of RT primer (AACAGCATCAGGAGTGGACAGATCCC) and ddH_2O to a final volume of 6.5 uL. Primer annealing was performed in a thermo-cycler set at 65ºC for 5 min, at which point the samples were incubated on ice for 1-2 min. To complete the setup of the RT PCR reaction, 2 uL of 5x RT buffer, 0.05 µmol DTT, 20 units of RNase (RNase-out ™, Invitrogen) and 100 units of reverse transcriptase (Superscript III Reverse Transcriptase, Invitrogen) was added per reaction, to a final volume of 10 uL. The RT reaction was completed in a thermo-cycler set to the following program: 45' at 50ºC, 15' min at 70ºC followed by incubation of the cDNA samples on ice. PCR amplification of the cDNA samples for the mini-gene splicing assay (3.10) was performed as follows: 6 uL of cDNA sample was mixed with 10 pmol of forward and reverse DUP51 primers (Forward: GACACCATCCAAGGTGCAC, Reverse: CTCAAAGAACCTCTGGGTCCAAG), respectively, 2.5 ul of 10x Taq polymerase buffer (Invitrogen), 37.5 nmol Mg^{2+} (Invitrogen), 5 nmol dNTPs and 2.5 units of Taq polymerase (NEB), to a final volume of 25 ul, and amplified in a thermocycler using the following program:
Template denaturation at 94ºC for 3 min, 15 cycles: 94ºC for 30s (DNA denaturation), 60ºC for 30s (primer annealing), 72ºC for 30s (primer extension). A final extension step for 3 min at 72ºC is added after completion of the last cycle. The PCR-amplified cDNA samples were run on a 2% agarose gel stained with Serva DNA Stain G or Ethidium Bromide.

Chapter 4. Results

4.1 Study of IQGAP1-hnRNPM RBP complex

4.1.1 IQGAP1 enters the nucleus and interacts with hnRNPM

Initially, cellular fractionation was performed for both cell lines of interest, and the nucleo-cytoplasmic distribution of IQGAP1 was examined by western blotting in nuclear extracts, in comparison to cytoplasmic ones. This analysis confirmed the nuclear localization of IQGAP1, while also verifying the presence of hnRNPM in the same nuclear fractions (Figure 15A). To ensure that IQGAP1 detection in nuclear extracts was not due to cross-contamination during the process of nuclear fractionation, probing with GAPDH and Lamin B1 antibodies was used as a cytoplasmic and nuclear control, respectively. The greatly reduced quantity of GAPDH in the immunoblot of nuclear extracts (Figure 15A) indicates absence of contamination from the cytoplasmic fraction, therefore validating that the presence of IQGAP1 in the nuclear extract is a result of the protein's nuclear translocation. Cellular fractionation followed by protein co-immunoprecipitation from nuclear extracts was performed for both cell lines, using antibodies for both IQGAP1 and hnRNPM. Nuclear extracts immunoprecipitated with normal rabbit and mouse IgG antibodies were used as negative controls (Figure 15B). Immunoblotting of anti-IQGAP1 IP samples showed co-immunoprecipitation of IQGAP1 with similar amounts of hnRNPM from both cell lines (Figure 15C). The interaction was validated by reciprocal co-immunoprecipitation of nuclear IQGAP1 with anti-hnRNPM antibody (Figure 15D). These results confirmed that IQGAP1 enters the nucleus where it interacts with hnRNPM, with what the lab's results have shown to be an RNA-independent protein-protein interaction.

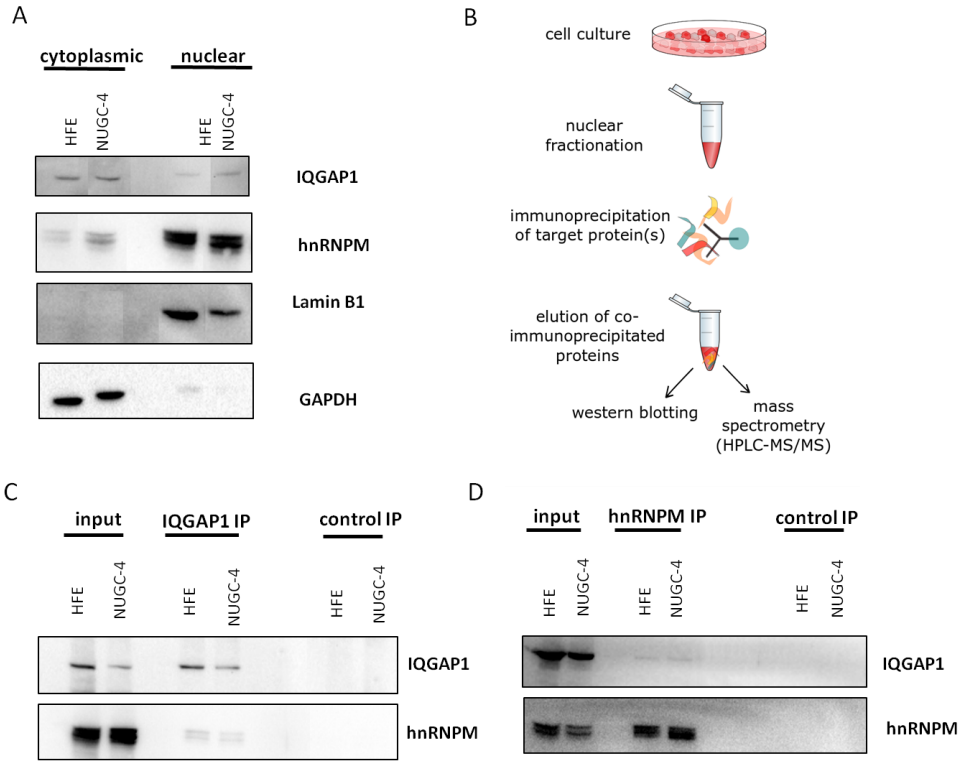

Figure 15. IQGAP1 interacts with hnRNPM. A. Experimental workflow of IQGAP1 immunoprecipitation in nuclear extracts of HFE and NUGC-4 cell lines. B. Immunoblotting of cytoplasmic and nuclear extracts from HFE and NUGC-4 cell lines. GAPDH and Lamin B1 were used as loading controls. C. Immunoblotting of IQGAP1 IP samples using anti-IQGAP1 rabbit polyclonal antibody. Input is 1/25th of the amount of protein used for the IP. ; control IP: rabbit normal IgG. D. Immunoblotting of hnRNPM IP samples using anti-hnRNPM mouse monoclonal antibody). Input is 1/20th of the amount of protein used for the IP; control IP: normal mouse IgG.

4.1.2 Nuclear IQGAP1 interacts with spliceosomal components

4.1.2.1 Proteomics analysis of IQGAP1 nuclear interactors in normal and gastric cancer cell lines

In order to shed light on the function of nuclear IQGAP1 and its potential role in splicing regulation, we first needed to identify its unknown protein interactors. The identification and quantification of proteins utilizing mass spectrometry-based proteomics techniques is a widely used strategy for the elucidation of protein composition as well as differential expression and enrichment of identified proteins in a variety of samples.[9] In the present study, proteins co-immunoprecipitated with IQGAP1 from nuclear extracts of HFE-145 and NUGC-4 cell lines (Figure 15C) were processed for HPLC-MS/MS analysis using standard processing techniques (see 3.5). In order to aid the comparative quantification of identified proteins with increased specificity, an isobaric labeling technique (Tandem Mass Tags ™)[262] was used (see 3.5). The analysis was performed in three biological replicates for each cell type (normal = HFE-145, cancer = NUGC-4). In addition to IQGAP1 co-immunoprecipitated samples, extracts from both cell lines were immunoprecipitated with normal rabbit IgG antibody, to serve as negative controls (HFE-145 IgG IP= normal IgG, NUGC-4 IgG IP= cancer IgG). Quantitative comparisons were performed between proteins identified in each IQGAP1 IP and respective control sample. The analysis identified a total 138 IQGAP1 interacting proteins enriched with robust statistical significance in IQGAP1 IPs compared to the rabbit polyclonal IgG used as isotype control. Quantitative comparisons were made between IQGAP1 IPs and their respective IgG isotype controls as well as between the HFE and NUGC-4 IQGAP1 IPs (cancer vs normal), using the LIMMA statistical analysis package.[264] Only proteins with more than 2 identified peptides were retained, and a protein was considered significant with a p-value < 5% (Benjamini-Hochberg FDR adjustment), and a fold-change of at least 50% between compared conditions. Enriched proteins were classified as hits, when showing a fold change > 2 between compared conditions, with a false discovery rate (FDR) lower than 0.05, and as candidates, when showing a fold change >1.5 and FDR < 0.25. The quantitative comparison between the HFE IQGAP1 IP and the respective HFE IgG control (IQGAP1 normal – IgG normal) yielded 135 enriched proteins (15 candidates, 120 hits) in HFE IQGAP1 IPs (Figure 16). The same comparison between IQGAP1 IP and IgG control in NUGC-4 cells (IQGAP1 cancer- IgG cancer) resulted in the identification of 120 enriched proteins (12 candidates, 108 hits, Figure 17).

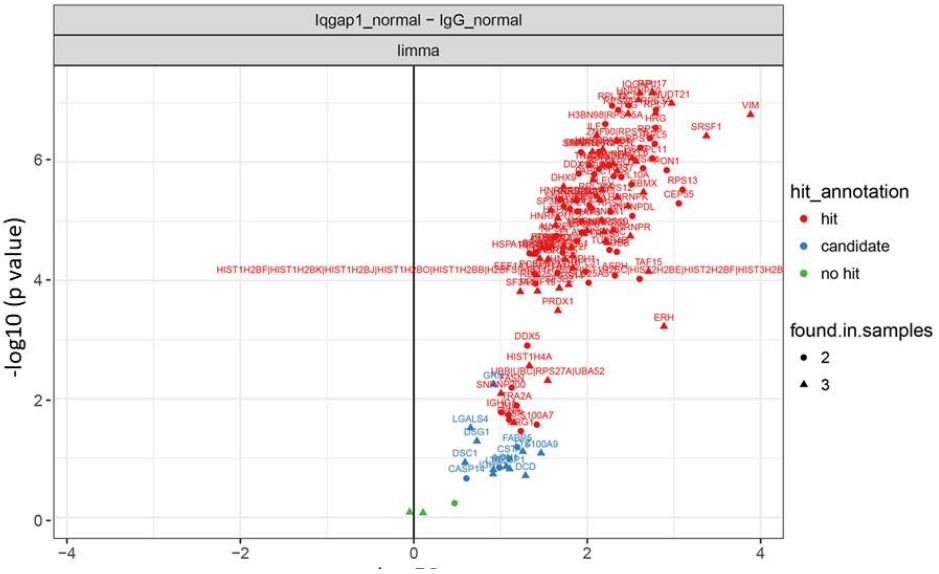

Figure 16. IQGAP1 interacting proteins in HFE-145 (normal) nuclear extracts.A volcano plot of quantitative comparison of proteins in HFE (normal) IQGAP1 IPs against their respective IgG controls, is shown. The x-axis indicates fold change between compared conditions and the y-axis the respective negative log10 (p value). Enriched proteins are classified into three categories: Hits (FDR threshold= 0.05, fold change=2)- denoted in red, candidates (FDR threshold = 0.25, fold change = 1.5)- denoted in blue and no hits-denoted in green.

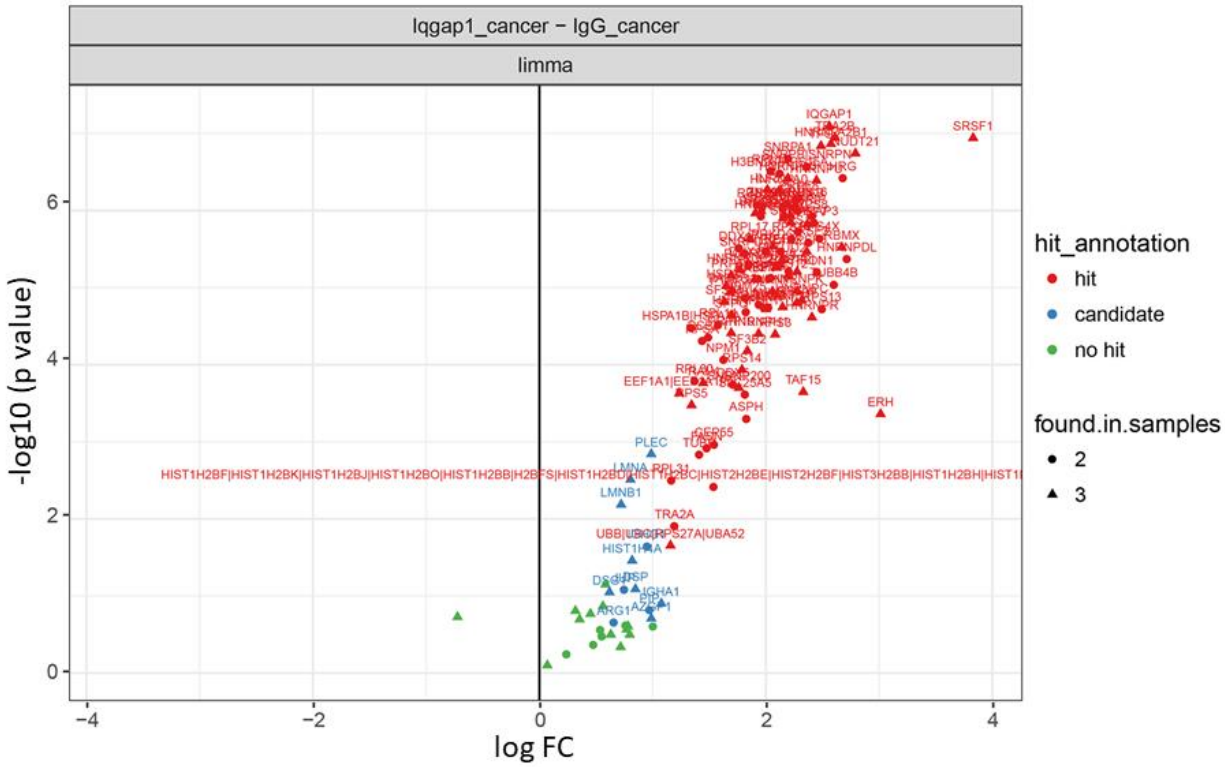

Figure 17. IQGAP1 interacting proteins in NUGC-4 (cancer) nuclear extracts. A volcano plot of quantitative comparison of proteins in NUGC-4 (cancer) IQGAP1 IPs against their respective IgG controls, is shown. The x-axis indicates fold change between compared conditions and the y-axis the respective negative log10 (p value). Enriched proteins are classified into three categories: Hits (FDR threshold= 0.05, fold change=2)- denoted in red, candidates (FDR threshold = 0.25, fold change = 1.5)-denoted in blue, and no hits- denoted in green.

However, quantitative comparison between the two sets of enriched proteins indicated that a robustly significant difference between the two conditions existed only in the enrichment of vimentin. (Figure 18). When comparing the enriched proteins between the HFE and NUGC-4 IPs non-quantitatively, we see that 15 proteins (including vimentin) are classified as hits or candidates in HFE-145 but not in NUGC-4 (Figure 19A), but a differential enrichment of those proteins is not supported by the quantitative comparison (Figure 18). In order to investigate the differential enrichment of vimentin, which showed a 9.84 fold-change enrichment in normal compared to cancer cells (logFC= -3.29), we examined the distribution of vimentin in HFE and NUGC-4 nuclear extracts and their respective IQGAP1 IPs. (Figure 19B) It is obvious that vimentin is present in HFE cytoplasmic and nuclear extracts but absent in NUGC-4. Immunoblotting with GAPDH in the same cytoplasmic and nuclear extracts (Figure 19B) shows a significant presence of GAPDH in inputs, indicating a contamination with the cytoplasmic fraction, however, the results of vimentin immunoblotting shown in the same figure also indicate that the difference in vimentin immunoprecipitation reflects a quantitative difference in vimentin expression between the cell lines of interest and not a differential interaction of IQGAP1 with vimentin between cell lines. As a result, the role of vimentin was not investigated further during the study presented here.

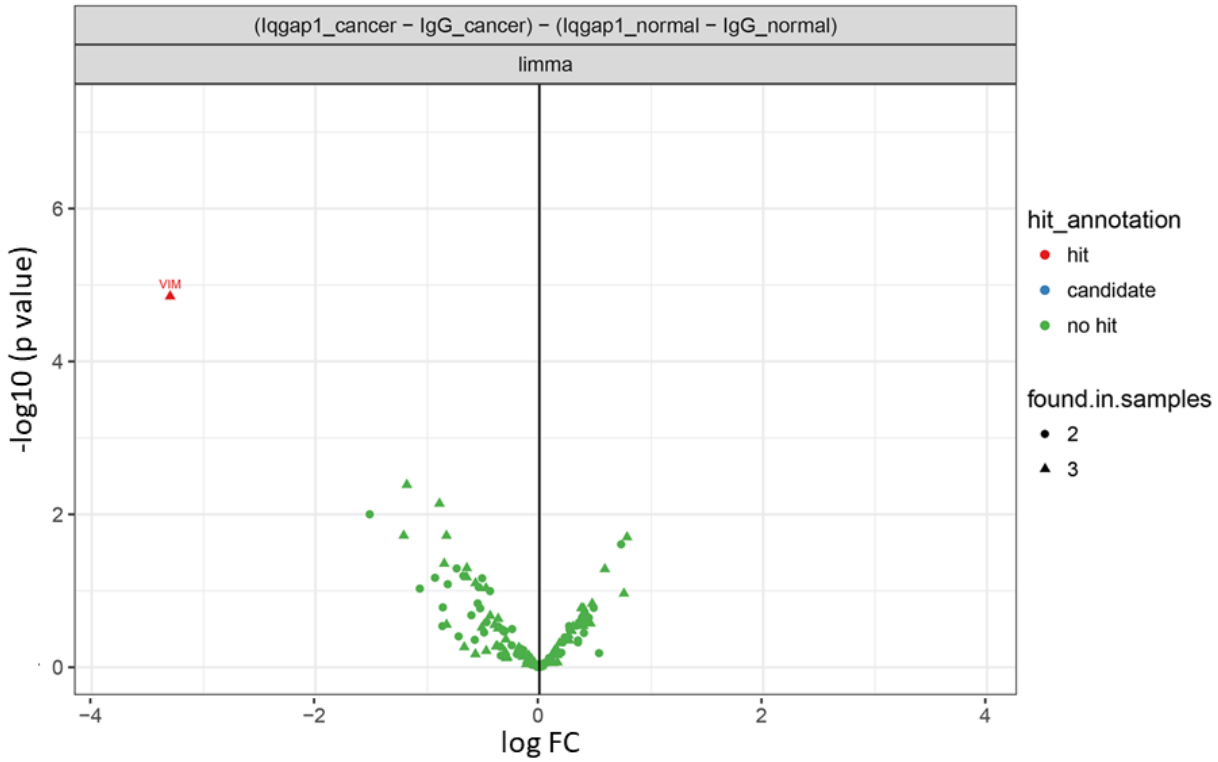

Figure 18. Quantitative comparison of IQGAP1 IP enriched proteins between HFE-145 (normal) and NUGC-4 (cancer). A volcano plot of quantitative comparison of protein enrichment between both sets of hits is shown. The x-axis indicates fold change between compared conditions and the y-axis the respective negative log10 (p value).. Enriched proteins are classified into three categories: Hits (FDR threshold= 0.05, fold change=2), denoted with red, candidates (FDR threshold = 0.25, fold change = 1.5), denoted with blue, and no hits, denoted with green.

Since the quantitative comparisons did not indicate any other statistically significant differential enrichment between the normal and gastric cancer cell lines, the analysis was focused on the full set of 138 nuclear IQGAP1 interactors identified through the proteomics analysis.

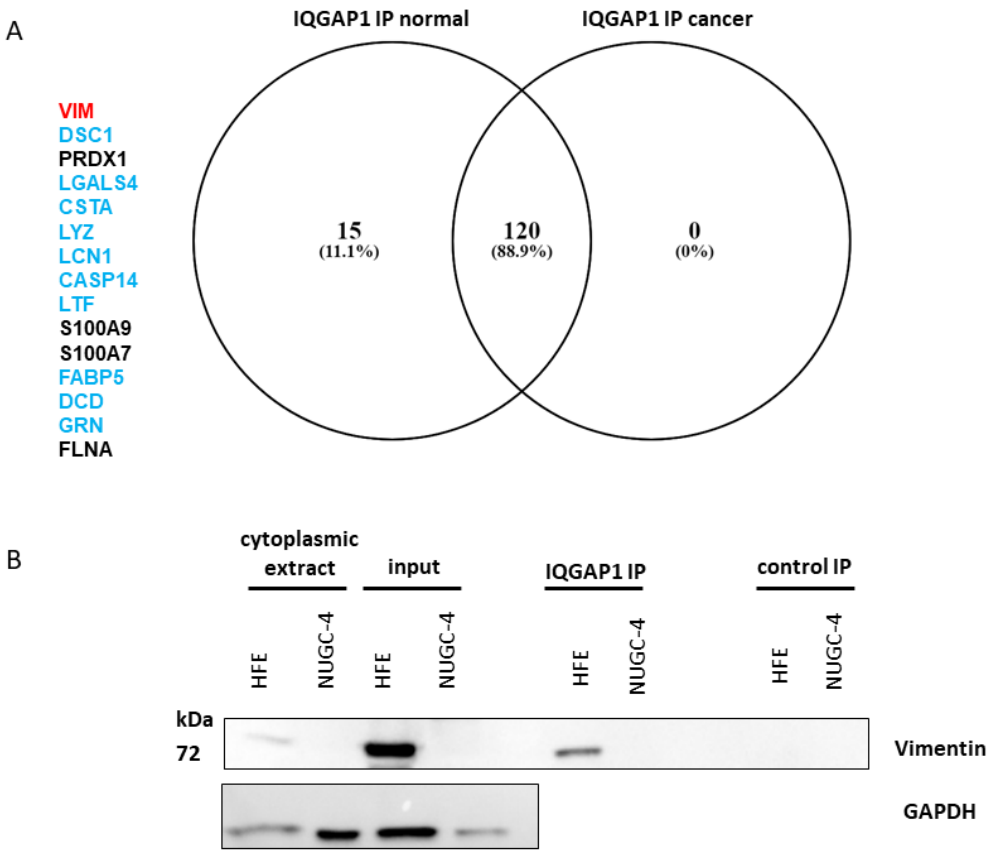

Figure 19. Analysis of differentially enriched proteins in HFE-145 IQGAP1 IPs vs NUGC-4 IQGAP1 IPs. A. Venn diagram comparing hits/candidates between IQGAP1 IPs in cancer (120 proteins) and normal cells (135 proteins). 15 proteins that were classified as hits (black) or candidates (blue) in the IP vs IgG control comparison in HFE-145 cells were classified as "no-hit" in the same comparison for NUGC-4, with the exception of vimentin (red). B Representative immunoblots of the cytoplasmic extracts, nuclear extracts (inputs), IQGAP1 and rabbit IgG control IP samples used for mass spectrometry analysis. IP samples: IQGAP1 IP (anti-IQGAP1 rabbit polyclonal antibody) and control IP (rabbit normal IgG); Input is 1/25th of total protein used in each IP sample. For the immunoblot, antibodies against vimentin and GAPDH (loading control) were used.

4.1.2.2 Gene ontology analysis of identified nuclear IQGAP1 interactors

Analysis of modular enrichment of gene ontology terms (Cellular Component-CC, Biological Process-BP and Molecular Function-MF) [275–277] showed that the interactors of nuclear IQGAP1 are greatly enriched in spliceosomal components and RNA binding proteins (Table 1). Significantly, 39 of the identified IQGAP1 interactors showed multiple annotations for nucleus (CC), nuclear mRNA splicing via spliceosome (BP) and RNA splicing (BP).

Table 1. Modular enrichment analysis of 135 nuclear IQGAP1 interactors (for all gene ontology annotations) performed in Genecodis online platform. The top 10 enriched annotation groups are shown, shorted by ascending corrected hypergeometric p value. NG= number of annotated genes in submitted gene list, NGR= number of annotated genes in reference gene list (Homo sapiens).

Enriched terms	Annotations	NG	NGR	Hyp_c p value
GO:0010467	**gene expression (BP)**	60	408	6.31832E-79
GO:0003723,GO:0010467	**RNA binding (MF),gene expression (BP)**	44	139	1.8933E-72
GO:0003723	**RNA binding (MF)**	58	623	2.24306E-64
GO:0005634,GO:0000398	**nucleus (CC),nuclear mRNA splicing, via spliceosome (BP)**	40	160	6.40347E-61
GO:0005634,GO:0000398, GO:0008380	**nucleus (CC),nuclear mRNA splicing, via spliceosome (BP),RNA splicing (BP)**	39	153	9.32432E-60
GO:0005634,GO:0008380	**nucleus (CC),RNA splicing (BP)**	42	260	1.35601E-55
GO:0005634,GO:0005654, GO:0000398,GO:0008380	**mRNA splicing, via spliceosome (BP),RNA splicing (BP)**	34	113	1.38516E-54
GO:0005634,GO:0005654, GO:0008380	**nucleus (CC),nucleoplasm (CC),RNA splicing (BP)**	35	132	4.06933E-54
GO:0005634,GO:0000398, GO:0008380,GO:0010467	**nucleus (CC),nuclear mRNA splicing, via spliceosome (BP),RNA splicing (BP),gene expression (BP)**	33	106	1.43361E-53
GO:0005515,GO:0003723, GO:0010467	**protein binding (MF),RNA binding (MF),gene expression (BP)**	31	81	1.46371E-53

The most enriched gene ontology[278,279] annotations for cellular component (CC) were: ribonucleoprotein complex ((p value = $5.78x10^{-69}$, 78 genes), spliceosomal complex (p value = $8.28x10^{-38}$, 35 genes) and catalytic step 2 spliceosome (p value =$3.82x10^{-37}$, 29 genes) (Figure 20A). Molecular function was enriched for RNA binding, nucleic acid binding and mRNA binding, with 105, 108 and 35 genes, respectively (Figure 20B), while the top three annotations for biological process were mRNA metabolic process, mRNA processing and mRNA splicing, via spliceosome (Figure 20C). Reflecting this clear enrichment in spliceosomal components, the KEGG pathways identified as significantly enriched were: ribosome, spliceosome, mRNA surveillance pathway, arythmogenic ventricular cardiomyopathy and RNA transport (Figure 20D). Interestingly, out of the total 31 proteins which were identified as ribosomal components, 29 were annotated as components of the cytosolic ribosome. However, closer inspection of these components showed that 20 of the ribosomal proteins were also annotated as having nuclear/nucleolar localization, while all identified ribosomal components were annotated (in Reactome) for participation in nuclear processes such as ribosomal biogenesis, rRNA processing, or nonsense-mediated decay (NMD). A number of these have been previously implicated in mRNA splicing regulation/spliceosomal assembly: Comparison of our results with spliceosomal components identified in previous mass spectrometry studies of the spliceosome, as compiled in the spliceosome database[280], showed that 71 of the identified IQGAP1 nuclear partners have

been previously annotated as components of spliceosomal complexes in mass spectrometry experiments (51.4% of the total 135 identified IQGAP1 interacting proteins). Both this comparison as well as an analysis of our enriched set through the Reactome database[281] did not show an enrichment for any specific spliceosomal complex. However, these results strongly support the hypothesis of IQGAP1's involvement in the regulation of post-translational modifications and/or assembly of spliceosomal components.

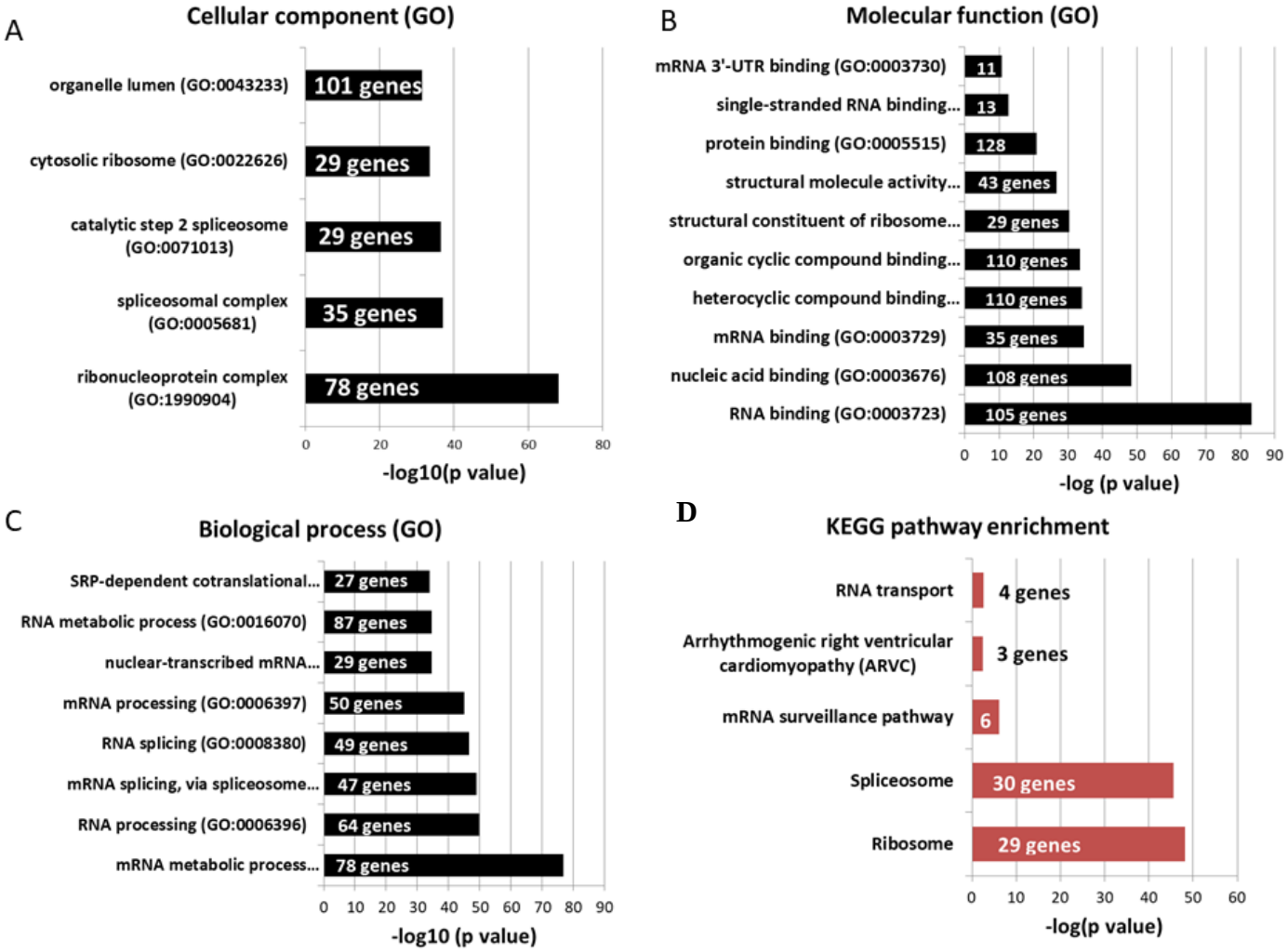

Figure 20. Gene Ontology and KEGG pathway enrichment analysis for the 138 IQGAP1 nuclear interactors. The analysis was performed using the GeneOntology and Genecodis online platforms for GO terms and KEGG pathways, respectively. A. 5 top enriched cellular component annotations. B. 10 top enriched molecular function annotations. C. Top 10 enriched biological process annotations. D. Enrichment of KEGG pathways.

4.1.3 IQGAP1 forms two putative complexes with spliceosomal components

After annotation of our putative complex components, we proceeded to further investigate these interactions (Figure 21). We selected for validation hits or candidates which were highly enriched in IQGAP1 IPs (in normal or cancer cells). These included hnRNPM, which had been previously validated (Figure 15C) and SRSF1 (ASF/SF2), which showed the second and first highest fold change compared to control in HFE-145 and NUGC-4 respectively (Figure 21A). In addition to these, other splicing factors and hnRNPs emerged as highly enriched candidates and their interaction with IQGAP1 was confirmed through immunoblotting. These included hnRNPL, hnRNPA1 and hnRNPC1/2 (Figure 21A) as well as DDX17 (Table 2). Because of SRSF1's prominent role in splicing regulation, as well as its apparently strong interaction with IQGAP1, a reciprocal IP using an anti-SRSF1 antibody was performed (Figure 21B). In an effort to elucidate the dynamic interactions of IQGAP1, splicing factors identified and validated in IQGAP1 IPs were probed in SRSF1 as well as hnRNPM IPs (Figure 21B). This experiment confirmed SRSF1 nuclear interaction with IQGAP1 in both cell lines, as we could observe immunoblotting of IQGAP1 in SRSF1 IP samples (Figure 21B). However, SRSF1's interaction with IQGAP1 is not consistently validated via immunoblotting of IQGAP1 IPs (data not shown). As a result, our current hypothesis, which will have to be examined further, is that IQGAP1-SRSF1 interaction is a transient, RNA-dependent interaction. Probing of additional splicing factors confirmed interactions between hnRNPA2B1 and the DHX9 helicase, which are co-immunoprecipitated with IQGAP1 (Figure 21A and data not shown) with both SRSF1 and hnRNPM. Interestingly, however, there doesn't appear to be an interaction between SRSF1 and hnRNPM, even though both interact with IQGAP1 (Figure 21B).

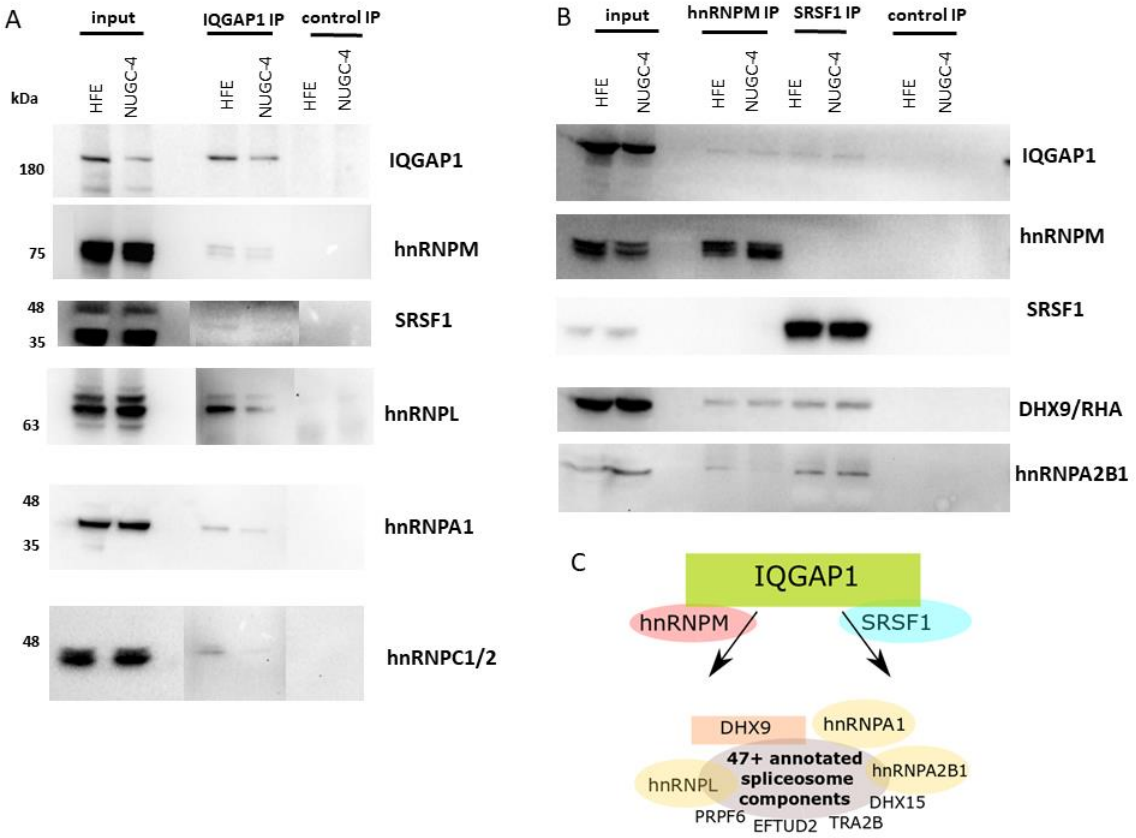

Figure 21. Investigation of the composition of IQGAP1 nuclear RBP complex. A. Validation of IQGAP1 nuclear interactions with various hnRNPs, via immunoblotting of IQGAP1 IP samples (anti-IQGAP1 rabbit polyclonal) with antibodies specific for the proteins noted at the right of each gel. B. Immunoblotting of hnRNPM and SRSF1 IP samples (anti-hnRNPM and anti-SRSF1 mouse monoclonal antibodies)control IP: mouse normal IgG. C. Schematic of the proposed IQGAP1 nuclear complexes, based on the investigation of mass spectrometry identified spliceosomal components. IQGAP1 is shown to form two putative complexes with hnRNPM and SRSF1 as key components, both of which are comprised of identified spliceosomal components (proteins validated by western blot are coloured).

This result, combined with the known primary functions of hnRNPM and SRSF1 as an alternative splicing silencer and enhancer, respectively, led us to the tentative hypothesis that IQGAP1 could regulate the assembly of two different (or more) spliceosomal or splicing-related complexes (Figure 21C). This hypothesis remains to be confirmed through ongoing further study of the protein interactions involved, some of which appear to be RNA-dependent (Table 2, data not shown).

Table 2. Summary table of western blot validations of RNA-dependent or independent interactions as seen by co-immunoprecipitation experiments. RNA dependency was examined by adding RNase to the nuclear lysate prior to immunoprecipitation, in which case RNA-dependent interactions are no longer detected by western blot. +/-: a pool of the interaction is not RNA-dependent; n.t.: not tested

Protein	**IP**			**Interaction with IQGAP1**
	anti-IQGAP1	**anti-hnRNPM**	**anti-SRSF1**	**RNA-dependent**
IQGAP1	+	+	+	
hnRNPM	+	+	-	-
SRSF1	+	-	+	+
hnRNPL	+	+	+	+/-
hnRNPA1	+	+	+	+
hnRNPC1/C2	+	+	+	+/-
hnRNPA2B1	+	+	+	+
DHX9	+	+	+	+
DDX17	+	**n.t**	**n.t**	

4.1.4 Study of post-translational modifications in IQGAP1-RBP complex components

The next step after the identification of the components of IQGAP1's nuclear RBP complex was to study the potential effect of IQGAP1 on alternative splicing regulation through the induction of post-translational modifications on its spliceosomal partners. Such a function would be expected, given IQGAP1's role as a scaffold protein, participating in numerous signaling pathways. In order to investigate post-translational modifications, we performed IQGAP1 co-immunoprecipitation on nuclear extracts of HFE-145 and NUGC-4, as described in Materials and Methods (3.6). Sample preparation was performed using two methods: Filter-Assisted Sample Preparation (FASP)[261], and an adapted form of the SP3 sample preparation strategy[265], in an effort to increase peptide yield (see section 3.6). The SP3 technique, as used in this experiment, was unsuccessful in obtaining a higher protein yield/ increased sensitivity of identification compared to the widely used FASP methodology, as the protocol used by the BSRC "Alexander Fleming" Proteomics Facility is still under optimization. Peptides were screened through a human phosphopeptide library compiled by the facility. Two of the previously identified IQGAP1 interacting proteins (HSPA8 and PRPF6) were shown to be phosphorylated. This is an interesting finding, since the function of tri-snRNP component PRPF6- which has been shown to promote cell proliferation in colon cancer through alternative splicing [173]- is highly dependent on its phosphorylation[131,282] (see Discussion).

Additional phosphoproteins were also identified, including 8 proteins annotated for RNA metabolism and/or mRNA splicing in the Reactome database[281]: PC4 and SRSF1-interacting protein (PSIP1), cleavage and polyadenylation specificity factor subunit 4 (NEB1), Telomerase-binding protein EST1A, CUGBP Elav-like family member 4 (CELF4), A-kinase anchor protein 17A (AKAP17A), GC-rich sequence DNA-binding factor 2 (GCFC2), Eukaryotic translation initiation factor 2A (EIF2A), Queuine tRNA-ribosyltransferase catalytic subunit 1 (QTRT1), Probable rRNA-processing protein EBP2 (EBNA1BP2) and H/ACA ribonucleoprotein complex subunit 1 (GAR1). An additional IQGAP1 interacting phosphoprotein which presents interest is Nuclear fragile X mental retardation-interacting protein 2 (NUFIP2). NUFIP2 is a shuttling RNA binding protein which interacts with FMRP[283], the RBP responsible for the Fragile X syndrome, whose pathogenesis is associated with alternative splicing of FMRP itself as well as alternative splicing of its interacting partners.[284,285] Other phosphoproteins detected include signaling pathway components, such as MAPK8 or CERS5 (Ceramide Synthase 5) which is responsible for the synthesis of C16 ceramides,[286] signaling molecules with an anti-apoptotic role. [287]

A more expansive study could validate these results and possibly uncover more targets for investigation, or even differential profiles between normal and adenocarcinoma cells (see Discussion).

4.2 Phenotypical study of IQGAP1 knockout cell lines

4.2.1 Generation of IQGAP1 knockout cell lines

In order to investigate the potential effect of IQGAP1 on the cancer phenotype of gastric cancer cells, through effects on alternative splicing, IQGAP1 knockout cell lines were generated in the lab, using the CRISPR-Cas9 strategy.[288] The two synthetic guide RNAs used to target exon 1 of the IQGAP1 gene were designed using the Optimized CRISPR Design Tool (Zhang Lab, MIT) [268], and targeting two adjoining regions, both bearing a PAM sequence (Figure 22A). The sgRNAs were cloned in All-in-one Cas9D10 nickase vector[269] containing both the Cas9-D10 nickase gene as well a puromycin resistance gene. Successfully transfected clones were initially selected for antibiotic resistance and screened with PCR. The clones showing depletion of IQGAP1 in PCR screening were screened for IQGAP1 expression with immunoblotting (Figure 22B). At least two IQGAP1 knockout clones were successfully generated from each parent cell line (HFE-145, NUGC-4), showing depletion of IQGAP1 expression. These were designated HQ15 and HQ17/2 (for HFE-145) and NQ16/3, NQ47/1, NQ47/3 (for NUGC-4). Additional clones (namely HQ17, NQ47) showed a knock-down of IQGAP1 expression compared to controls (Figure 22C). For the experiments described below, we used clones HQ15, HQ17/2, which we will refer to as HQ and NQ16/3, NQ47/1, which we will refer to as NQ.

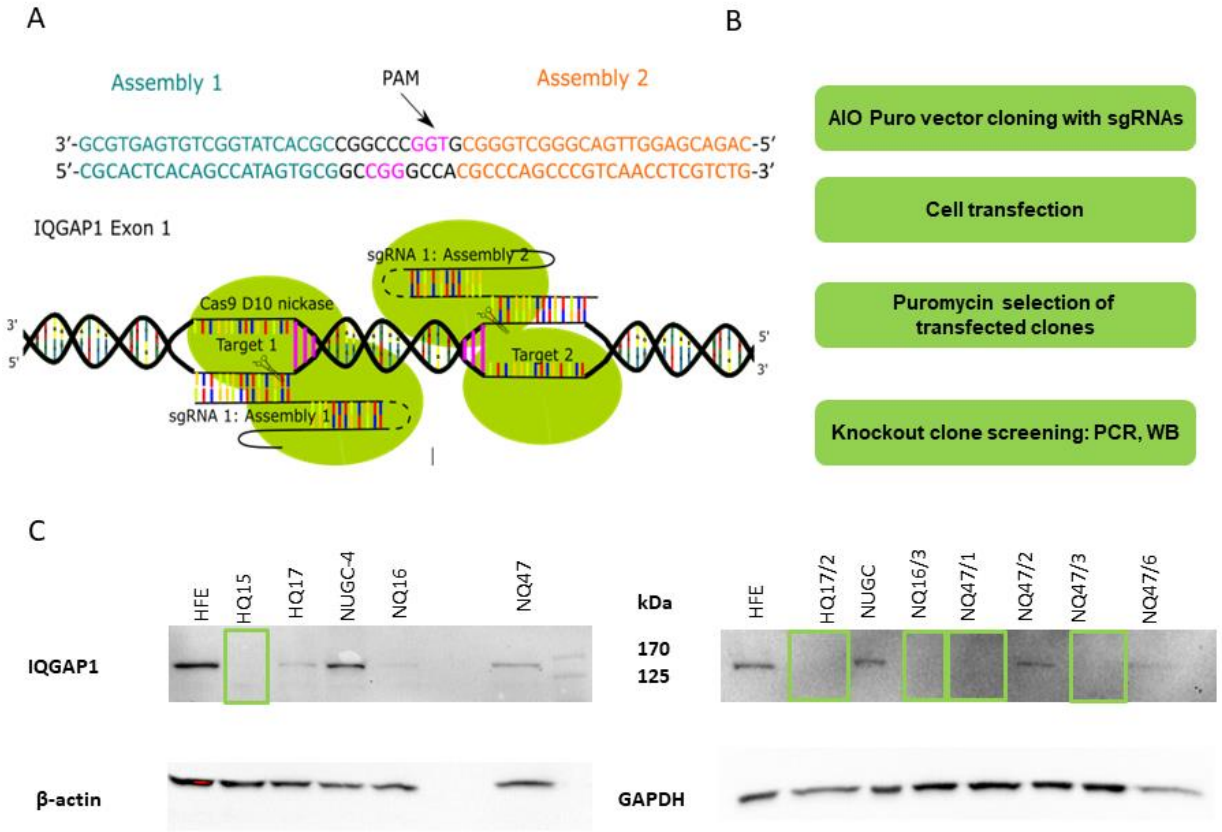

Figure 22. Generation of IQGAP1 knockout cell lines. A. Two synthetic sgRNA assemblies (Assembly 1 and Assembly 2) targeting adjoining regions of exon 1 of the IQGAP1 gene were cloned in a vector (AIO Puro) containing the Cas9-D10 nickase. B. Experimental strategy for the generation of knockout clones. C. Representative immunoblots of puromycin selected clones, checking the depletion of IQGAP1 expression (anti-IQGAP1 primary antibody). β-actin and GAPDH were probed as loading controls. Clones showing successful depletion of IQGAP1 expression are indicated in green rectangles.

4.2.2 Effects of IQGAP1 depletion on cell cycle progression on normal gastric epithelial and gastric cancer cells

In order to study the effect of IQGAP1 knockout on the phenotype of cancer cells, we employed strategies which are widely used to study characteristics of cancer, including adhesiveness, invasion, proliferative potential and cell cycle progression.[289] The effect of IQGAP1 depletion on cell cycle progression was investigated with fluorescence assisted cell sorting of cells stained with propidium iodide, which stains DNA, separating the cell cycle phases according to their DNA content (see 3.9) The CRISPR-Cas9 D10 IQGAP1 depleted clones of the HFE-145 and NUGC-4 cell lines were cultured to a 55-75% confluence, in unsynchronized conditions, and stained with propidium iodide after ethanol fixation and washing with phosphate-citrate buffer. RNAse was added to the staining step to ensure only DNA was stained. Analysis was performed in three independent experiments for IQGAP1 depleted clones and cells from the respective parental cell lines were used as controls (Figure 23).

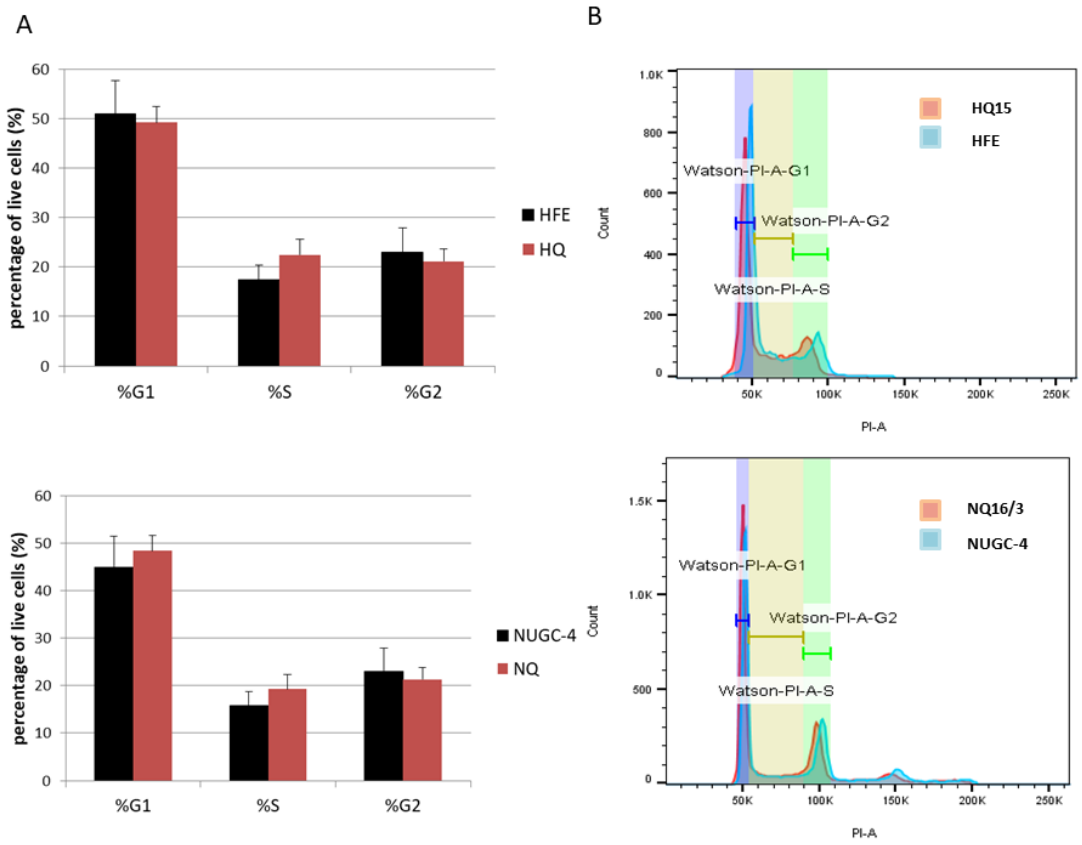

Figure 23. Cell cycle progression in IQGAP1 knockout cells. A. Percentage of live cells with DNA content corresponding to each phase of the cell cycle (G1/S/G2-M). HFE, NUGC-4: n= 3 technical replicates, HQ: n=4 technical/biological replicates NQ: n = 4 technical/biological replicates. Statistical significance measured by unpaired, two tailed, equal variances student's t test. Error bars represent standard deviation. B. Representative histograms of PI stained live populations: HQ15 (blue) is overlayed to HFE-145 (red) (top) and NQ16-3 is overlayed with NUGC-4 (bottom).

The FACS results indicate that IQGAP1 depletion has no significant effect on the cell cycle progression of normal gastric epithelial HFE-145 cells or NUGC-4 gastric adenocarcinoma cells (Figure 23A). Side by side comparisons of PI-staining histograms between control and IQGAP1 depleted cells highlight the fact that the number of cells in each phase of the cell cycle is similar between wild type and IQGAP1 depleted clones (Figure 23B). These results concur with what was previously observed in IQGAP1-siRNA silenced NIH3T3 cells, which also did not display a differential cell cycle progression after asynchronous culture. [290] (see Discussion).

4.2.3 Effects of IQGAP1 depletion on proliferative capacity of normal gastric epithelial and gastric cancer cells

The proliferative capacity of the IQGAP1-depleted clones was assessed using a colony formation assay designed to test the clonogenic capacity of adhesive cells as described in 3.8.[270]

The survival fraction (SF) of each control was taken as 100%, and the SFs of the IQGAP1 depleted clones were compared to the control using a student's t test. The results shown in Figure 24 indicate that IQGAP1 depletion causes a decrease of about 50% in the clonogenic capacity of HFE-145 cells: we can see that HQ clones show a statistically significant decrease in survival fraction compared to the parental HFE cells. This result is in agreement with IQGAP1's known role in cell proliferation and adhesion. However, the reverse effect is observed in IQGAP1 depleted NUGC-4 cells: comparing the survival fraction of NQ clones with the parental NUGC-4 cells, we observed that the NQ cells exhibit an increased clonogenic capacity (Figure 24). Further experiments will need to be performed to verify this interesting result (see Discussion).

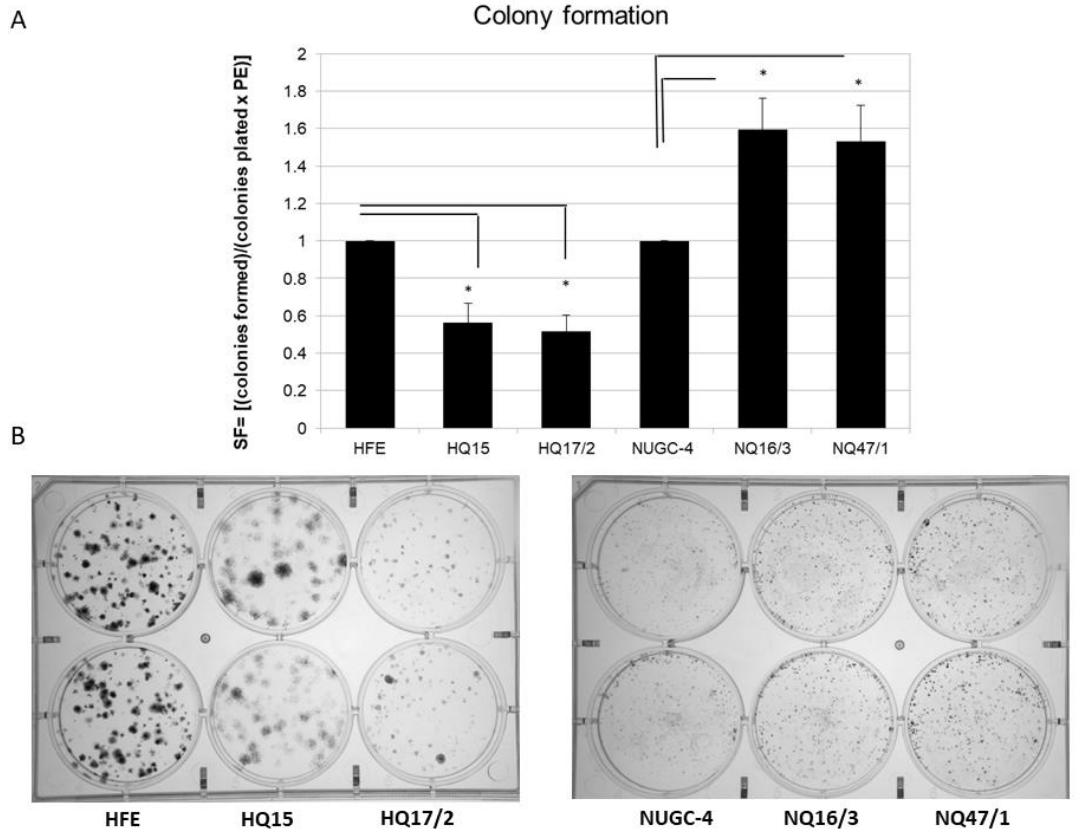

*Figure 24 Assessment of proliferative capacity of HQ and NQ cells. A. Calculation of survival fraction (SF) for IQGAP1 knockout clones. N= 3 independent experiments, survival fraction in controls normalized to 100%. Statistical significance *p value < 0.05, measured by unpaired two tailed student's t test with unequal variances. Error bars represent standard deviation. B. Representative image of colony formation in HFE-145, NUGC-4 and their respective knockout clones.*

4.2.4 Splicing effect of IQGAP1 knockout on hnRNPM-dependent splicing

Since the proteomics analysis of IQGAP1 immunoprecipitated samples in nuclear extracts showed that IQGAP1 interacts with a significant number of spliceosomal complex components, both in normal gastric epithelial cells as well as gastric adenocarcinoma cell line, we sought to investigate whether this interaction would have any observable effect on alternative splicing regulation. Because hnRNPM was the first observed spliceosomal component seen to interact with IQGAP1 in the nucleus, and one of the most robust and direct interactions observed in subsequent experiments, we utilized an assay targeted towards the investigation of hnRNPM-dependent splicing events. DUP51 is a three exon two intron construct developed for the study of differential exon inclusion,[274] which was subsequently adapted for the investigation of hnRNPM-dependent alternative splicing.[49] More specifically, exon 2 of the adapted DUP51M construct contains an hnRNPM consensus binding motif (UGGUGGUG), which acts as an exonic splicing silencer (Figure 25A).[273]

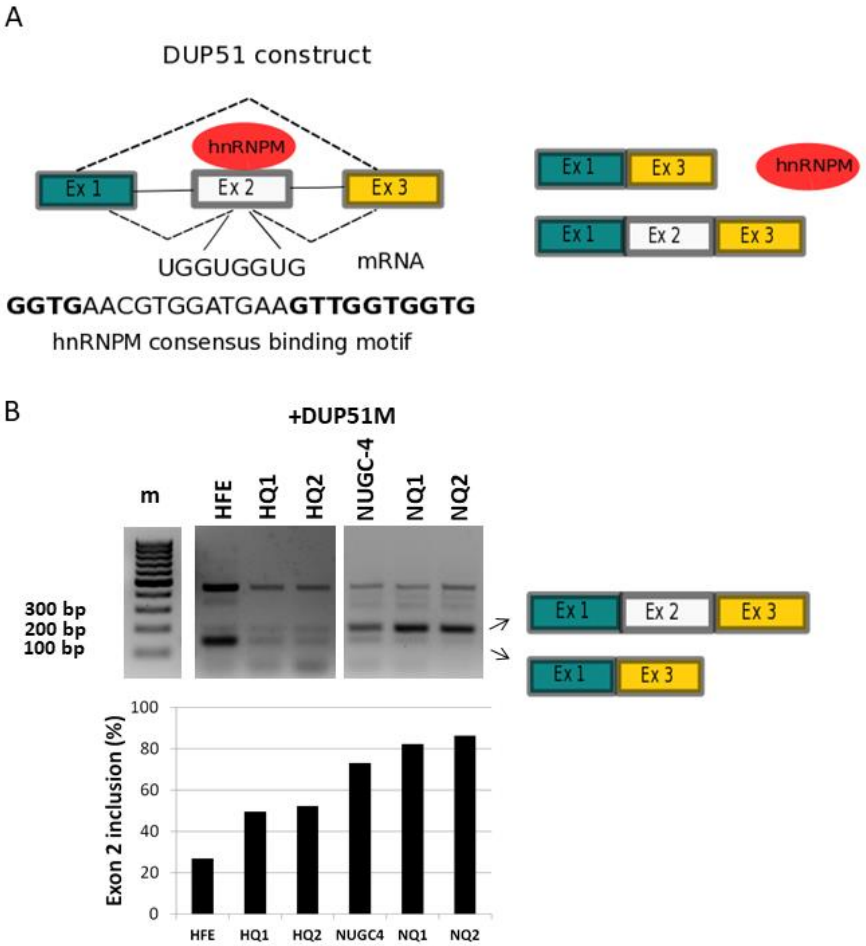

Figure 25 Splicing effect of IQGAP1 depletion on hnRNPM-dependent splicing, using the DUP51 three exon, two intron splicing construct. A. Diagram of hnRNPM functioning as an exonic splicing silencer. B. Representative splicing assay on RNA isolated from control (HFE, NUGC-4) and IQGAP1-depleted cells (HQ15, NQ16/3). Cells were transfected with the DUP51M construct and a control plasmid (4 µg total DNA per well). The two alternative splicing products of DUP51 pre-mRNAs_are indicated. Quantification of the splicing products was performed using ImageJ software. After quantification of band intensity, the ratio: [(Exon 1+3)/ (Exon1+2+3) +(Exon1+3)] x 100% was calculated for each lane.

When the splicing function of hnRNPM is unperturbed, its binding to the consensus motif of exon 2 leads to exon 2 of the DUP51 construct being spliced out, and the formation of the smaller, two-exon splicing product of construct DUP51. (Figure 25A) Using this strategy, the changes observed in the splicing product ratio between parental and IQGAP1-depleted clones appear to follow the same pattern (Figure 25 B): for both cell lines, IQGAP1 depletion results in less efficient repression of exon 2 by hnRNPM. Furthermore, we can observe that the two cell lines of interest exhibit a different splicing pattern, with NUGC-4 and NQ cells producing a higher ratio of exon 2 inclusion compared to HFE and HQ cells (Figure 25B), a result which will need to be verified by further experiments. In order to ascertain whether this differential pattern was due to a differential expression of hnRNPM between the parent cell lines, we examined hnRNPM distribution in HFE-145 and NUGC-4 nuclear extracts (Figure 15B and data not shown). However, no striking difference in hnRNPM expression was noted, therefore the mechanism of this effect remains to be elucidated.

Chapter 5. Discussion

The first objective of the present work was to identify the protein interactors of IQGAP1 in the nucleus and examine their function, in order to confirm whether or not IQGAP1 could play the role of an orchestrating scaffold in AS regulation. The results presented here confirm that IQGAP1 enters the nucleus of gastric epithelial and gastric cancer cells, where it interacts with numerous splicing factors and known spliceosomal components: 44 out of 135 proteins identified as IQGAP1 interacting proteins have RNA splicing as an annotated function, while a total of 72 out of the 135 of the proteins evidenced to co-immunoprecipitate with IQGAP1 have also been identified in previous mass spectrometry-based investigations of spliceosomal components. The results show that IQGAP1 engages in interactions with hnRNPs (including hnRNPM, hnRNPL, hnRNPA1, hnRNPA2/B1, hnRNPC1/C2 and hnRNPK/J), SRSFs (including SRSF1), while various other splicing factors, including UsnRNP-specific components were identified. No preference of interaction with IQGAP1 was shown in components belonging to any specific spliceosomal complex- IQGAP1's activity as an assembly point of splicing factors and spliceosomal components does not appear to play a role in the coordination of any particular step of the splicing reaction or mRNA processing. However, the fact that IQGAP1 appears to engage in distinct RNP complexes, interacting in an RNA-independent manner with hnRNPM and SRSF1 lead to us to conclude that IQGAP1 participates in the formation of at least two RNA-protein (RNP) complexes. Based on the known regulatory functions of hnRNPs and SRSFs, these IQGAP1 regulated complexes could have opposing roles in regulating the inclusion or excision of specific mRNA exons or introns in AS events of mRNA targets. On the other hand, this could be an indication that different mRNAs are targeted through IQGAP1 regulated snRNPs, depending on the complex composition and function- a question open to further investigation.

The second step of investigating the novel IQGAP1 RNP complex was to ascertain whether IQGAP1 also facilitates post-translational modifications of splicing factors through signal transduction. Our first preliminary investigation of phosphorylation patterns in IQGAP1 immunoprecipitated nuclear RNPs uncovered IQGAP1 interactions with phosphorylated spliceosomal components, including PRPF6, PRPF31 and hnRNPA2B1. Phosphorylation of these components is essential in their splicing function, and further experiments using the IQGAP1 knock-out cell lines generated in this lab would allow us to determine whether or not IQGAP1 is responsible for mediating these post-translational modifications. The low abundance of phosphorylated peptides in IQGAP1-immunoprecipitated nuclear RNPs has led to the development of an alternative experimental workflow, in which nuclear extracts of wild type and knock-out cell lines are processed with mass spectrometry, as described (see sections 3.2, 3.6), and the post-translational modifications of previously identified IQGAP1 interacting proteins are compared between IQGAP1 wild type and knock-out gastric epithelial and gastric cancer cell lines, in ongoing experiments. The evidence gathered by the present work suggests that IQGAP1 RNPs do not exhibit a differential composition between gastric cancer (NUGC-4) and normal gastric epithelial (HFE-145) cell lines. As a result, any role that the IQGAP1 RNPs may have in affecting tumor cell development and progression through AS regulation does not appear to be mediated through differential assembly of spliceosomal components. However, further study of post-translational modifications would help determine the role of IQGAP1 in modulating the function of splicing factors in both cancerous and physiological conditions.

To summarize, the first stage of this work helps expand the nuclear role of IQGAP1, which has already been implicated in RNA metabolism[205,291], by showing that nuclear IQGAP1 is an integral part of spliceosomal RNP formation in the nucleus, interacting with spliceosomal components participating in all stages of the splicing reaction. IQGAP1 appears to participate in at least two independent RNP complexes, which have

hnRNPM and SRSF1 as key components. A potential role of IQGAP1 in post-translational modifications of splicing factors and/or AS regulation of various RNA targets has begun to emerge, and will have to be explored with further experiments (Figure 26).

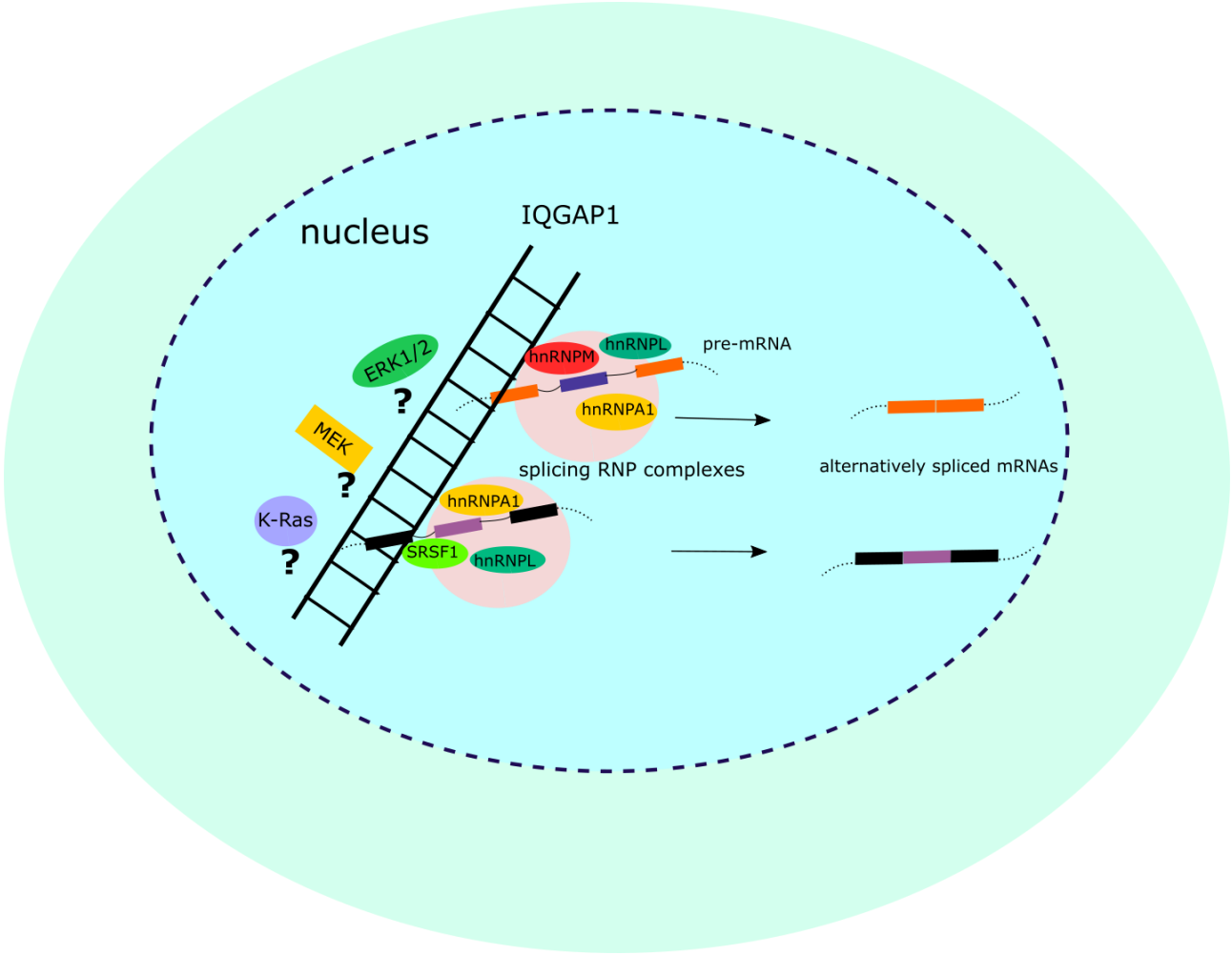

Figure 26. The proposed model for a nuclear role for IQGAP1 in splicing regulation. IQGAP1, which is a known binding partner of signaling pathway components such as ERK1/2, MEK and K-Ras, appears to act as a scaffold for at least two novel RNP complexes. The first is formed through IQGAP1's direct protein-protein interaction with hnRNPM as well as RNA-dependent interactions with other spliceosomal components, including hnRNPL and hnRNPA1. A second IQGAP1 RNP is formed through its interaction with SRSF1, as well as hnRNPL, hnRNPA1 and other protein components common between the two complexes.

The splicing assays using the successfully generated IQGAP1 knockout clones of NUGC-4 and HFE-145 cells support the hypothesis that IQGAP1-hnRNPM-mediated formation of RNP complexes is involved in AS regulation (see 4.2.4). Our results show that depletion of IQGAP1 quantifiably affects the hnRNPM-dependent AS pattern of the DUP51-M plasmid, by inhibiting hnRNPM mediated silencing of an alternatively spliced exon. This result needs to be verified, using a more sensitive experimental set-up (e.g. Radioactive RT-PCR), while add-back experiments need to be performed in order to verify rescue of the effect by exogenous IQGAP1. This experimental evidence indicates that IQGAP1 could play a potentially important role for the regulation of hnRNPM activity in pre-mRNA splicing *in vitro*.

The second objective of this study was to examine the effects of the depletion of IQGAP1 on the phenotype of normal gastric epithelial and gastric cancer cells, and explore whether any observed differences had a direct connection with processes regulated by AS.

The first phenotypical parameter we examined was cell cycle distribution, using PI-stained asynchronously cultured cells. The results show that IQGAP1 depletion appears to have no effect on the cycle progression of HFE-145 and NUGC-4 cells similarly to what has previously been observed in NIH3T3 cells[290]: Previous studies examining the effect of IQGAP1 on cell proliferation using the more sensitive BrdU incorporation assay in asynchronous culture had showed that IQGAP1 participates in the regulation of cell cycle progression through differential effects of its structural/functional domains.[292] The most definitive evidence of IQGAP1's role in cell cycle distribution comes from experiments using thymidine in order to induce arrest in the S-phase after IQGAP1 silencing with siRNA. These experiments showed that IQGAP1 depletion delayed G1/S phase progression- showing that IQGAP1 is required for G1/S phase progression. Most importantly, this study indicated that IQGAP1 undergoes slow nuclear transport and its nuclear accumulation is partially regulated by CRM1, and investigators proposed that IQGAP1 enters the nucleus at the G1/S phase border and exits in late S phase.[290] These results propose a role for IQGAP1 in cell cycle regulation which is related to IQGAP1 nuclear shuttling- obviously IQGAP1's interactions with spliceosomal machinery have not yet been implicated in this function. Our investigation of IQGAP1's role in cell cycle distribution and cell proliferation in HFE-145 and gastric adenocarcinoma cell lines had the goal of examining whether IQGAP1 was necessary in mediating cell cycle progression in normal or cancerous cells- cell cycle distribution and overall proliferation appear to be unaffected by IQGAP1 depletion in asynchronous culture. However, the nuclear shuttling of IQGAP1 through the CRM1 receptor between G1/S phase that has been observed by Johnson et al could be a starting point in order to definitively prove IQGAP1's role in splicing factor assembly: Studying AS patterns with RNA sequencing in IQGAP1 wild type cells could act in a similar manner to IQGAP1 over-expression- both techniques could help elucidate whether increased accumulation of IQGAP1 in the nucleus influences AS regulation.

The second phenotypical parameter examined in IQGAP1 knockout HFE-145 and NUGC-4 cells was clonogenic capacity, using a colony formation assay for adherent cells. Interestingly, IQGAP1 depletion seemed to have opposing effects on the ability of the two cell lines to form colonies in culture. HFE-145 clonogenic capacity decreased in IQGAP1-depleted cells compared to control, while the clonogenic capacity of NUGC-4 control cells appeared to be negatively influenced compared to their IQGAP1 depleted counterparts. This result might indicate a differential function for IQGAP1 in regulating cell adhesion or proliferation in the two cell lines examined- however further studies are necessary in order to confirm this effect. However, and interesting question is raised: Does IQGAP1 have a differential function in regulating cell adhesion and colony formation in gastric epithelial and metastatic cancer adenocarcinoma cells?
Future cell-based and in-vivo studies using orthotopic mouse models of gastric cancer invasion and metastasis could help shed light on the role of IQGAP1 and the IQGAP1/hnRNPM interaction in the development and progression of gastric adenocarcinoma.

To summarize the second part of this study, IQGAP1 appears to have a similar role in cell cycle progression and distribution in HFE-145 and NUGC-4 cells as that previously recorded in non-gastric cell lines- as a result, the physiological effects of IQGAP1 depletion on gastric tissue development don't appear to be mediated through changes in cell cycle progression- the gastric hyperplasia observed in IQGAP1 knock-out mice does not indicate that a negative effect of IQGAP1 depletion on cell proliferation should be expected. Interestingly, IQGAP1 depletion has opposing effects in regulating clonogenic capacity of the two cell lines examined, indicating a different role of IQGAP1 in regulating clone formation in HFE-145 gastric epithelial and NUGC-4 gastric adenocarcinoma cells. However further experiments will be required for the evaluation of this initial result. Finally, further studies using assays aimed at evaluating cell proliferation, adhesion,

migration and clonogenic capacity could be employed using both IQGAP1 depleted and hnRNPM/IQGAP1 double depleted cell lines in order to fully map the effect of this protein-protein interaction on the phenotype of cancer cells.

Continuing investigation will determine the impact of IQGAP1-containing RNPs in AS regulation in gastric cancer, and whether or not this AS regulation could be linked to gastric cancer phenotype. If one or both of these statements are true, the hnRNPM/IQGAP1 interaction could emerge as a potential target for therapeutic intervention in cancer.

Table of abbreviations

AS	Alternative Splicing
RNP	Ribonucleoprotein
RBP	RNA Binding Protein
snRNP	Small nuclear ribonucleoprotein
U-snRNA	Uridine-rich small nuclear RNAs
EJC	Exon-Junction Complex
TREX	Transcription-Export Complex
hnRNP	Heterogeneous nuclear Ribonucleoprotein
RBD	RNA-binding Domain
RRM	RNA Recognition Motif
SRSF	Serine/Arginine Rich Splicing Factor
3'/5' ss	3'/5' splice-site
3'/5' UTR	3'/5' Untranslated Region
SMD	Staufen-mediated decay
NMD	Nonsense-mediated decay
FASP	Filter-assisted sample preparation
HPLC	High Performance/Pressure Liquid Chromatography
CRISPR-Cas9	Clustered Regularly Interspaced Short Palindromic Repeats- CRISPR associated protein 9
sgRNA	Synthetic guide RNA
FC	Fold Change
FDR	False Discovery Rate
PI	Propidium Iodide

.

Addendum

Supplementary Table 1: Primary and secondary antibodies

Target	Use	Antibody	Concentration	Working concentration (WB)	supplier	Catalogue number
IQGAP1	WB, IP	Rabbit polyclonal	0.36 mg/mL	1:1000	Proteintech	22167-1-AP
IQGAP1	WB	Rabbit polyclonal	1.5 mg/mL	1:1000	Sigma-Aldrich	SAB4200329
IQGAP1 (C-9)	WB	Mouse Monoclonal	0.2 mg/mL	1:1000	Santa Cruz Biotechnology	Sc-376021
hnRNPM1-4 (1D8)	WB, IP	Mouse Monoclonal	0.2 mg/mL	1:1000	Santa Cruz Biotechnology	Sc-20002
SF2/ASF (96)	WB, IP	Mouse Monoclonal	0.2 mg/mL	1:1000	Santa Cruz Biotechnology	Sc-33652
Normal rabbit IgG	IP	Polyclonal Antibody	1 mg/mL		Millipore	12-370
Normal rabbit IgG	IP	Polyclonal Antibody	1.25 mg/mL		Proteintech	30000-0-AP
Normal mouse IgG	IP	Polyclonal Antibody	0.4 mg/mL		Santa Cruz Biotechnology	Sc-2025
Beta-actin	WB	Mouse Monoclonal	0.43 mg/mL	1:5000	Proteintech	66009-1-Ig
GAPDH	WB	Mouse Monoclonal	1 mg/mL	1:40.000	Proteintech	60004-1-Ig
Lamin B1 (A-11)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-377000
hnRNPA1 (4B10)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-32301
hnRNPK/J (3C2)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-32307
hnRNPL (4D11)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-32317
hnRNP C1/C2 (4F4)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-32308
IQGAP2 (4F4)	WB	Mouse Monoclonal	0.2 mg/mL	1:2000	Santa Cruz Biotechnology	Sc-17835
CEP55 (B8)	WB	Mouse Monoclonal	0.2 mg/mL	1:1000	Santa Cruz Biotechnology	Sc-374051
Vimentin (E5)	WB	Mouse Monoclonal	0.2 mg/mL	1:1000	Santa Cruz Biotechnology	Sc-373717
anti-rabbit IgG (H+L)	WB, HRP conjugated	goat, mouse and human adsorbed	1 mg/mL	1:2000	SouthernBiotech	4050-05
anti-mouse IgG	WB, HRP conjugated	goat, human adsorbed	1 mg/mL	1:2000	SouthernBiotech	1030-05

Bibliography

1. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).
2. Venter, J. C. *et al.* The Sequence of the Human Genome. *Science (80-.).* **291,** 1304–1351 (2001).
3. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science (80-.).* **306,** 636–640 (2004).
4. Human Genome Sequencing Consortium, I. Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–945 (2004).
5. Pertea, M. & Salzberg, S. L. Between a chicken and a grape: estimating the number of human genes. *Genome Biol.* **11,** 206 (2010).
6. Ezkurdia, I. *et al.* Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum. Mol. Genet.* **23,** 5866–78 (2014).
7. Yang, X. *et al.* Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. *Cell* **164,** 805–17 (2016).
8. Aebersold, R. *et al.* How many human proteoforms are there? *Nat. Chem. Biol.* **14,** 206–214 (2018).
9. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).
10. Roth, M. J. *et al.* Precise and parallel characterization of coding polymorphisms, alternative splicing, and modifications in human proteins by mass spectrometry. *Mol. Cell. Proteomics* **4,** 1002–8 (2005).
11. Kim, M.-S. *et al.* A draft map of the human proteome. *Nature* **509,** 575–81 (2014).
12. Ponomarenko, E. A. *et al.* The Size of the Human Proteome: The Width and Depth. *Int. J. Anal. Chem.* **2016,** (2016).
13. Smith, L. M., Kelleher, N. L. & Consortium for Top Down Proteomics, T. C. for T. D. Proteoform: a single term describing protein complexity. *Nat. Methods* **10,** 186–7 (2013).
14. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40,** 1413–1415 (2008).
15. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–6 (2008).
16. Nilsen, T. W. & Graveley, B. R. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463,** 457–463 (2010).
17. Kelemen, O. *et al.* Function of alternative splicing. *Gene* **514,** 1–30 (2013).
18. Liu, Y. *et al.* Impact of Alternative Splicing on the Human Proteome. *Cell Rep.* **20,** 1229–1241 (2017).
19. Blencowe, B. J. The Relationship between Alternative Splicing and Proteomic Complexity.

Trends Biochem. Sci. **42,** 407–408 (2017).

20. Black, D. L. Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* **72,** 291–336 (2003).

21. Lee, Y. & Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* **84,** 291–323 (2015).

22. Schmidt, V. & Kirschner, K. M. Alternative pre-mRNA splicing. *Acta Physiol.* **222,** e13053 (2018).

23. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc. Natl. Acad. Sci. U. S. A.* **75,** 4853–7 (1978).

24. De Conti, L., Baralle, M. & Buratti, E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* **4,** 49–60 (2013).

25. Choi, Y. D. & Dreyfuss, G. Isolation of the heterogeneous nuclear RNA-ribonucleoprotein complex (hnRNP): a unique supramolecular assembly. *Proc. Natl. Acad. Sci. U. S. A.* **81,** 7471–5 (1984).

26. DiMaria, P. R., Kaltwasser, G. & Goldenberg, C. J. Partial purification and properties of a pre-mRNA splicing activity. *J. Biol. Chem.* **260,** 1096–102 (1985).

27. Will, C. L. & Lührmann, R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **3,** (2011).

28. Chen, W. & Moore, M. J. The spliceosome: disorder and dynamics defined. *Curr. Opin. Struct. Biol.* **24,** 141–9 (2014).

29. Lerner, M. R. & Steitz, J. A. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl. Acad. Sci. U. S. A.* **76,** 5495–9 (1979).

30. Will, C. L. & Lührmann, R. Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell Biol.* **13,** 290–301 (2001).

31. Matera, A. G. & Wang, Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* **15,** 108–21 (2014).

32. Patel, A. A. & Steitz, J. A. Splicing double: insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* **4,** 960–970 (2003).

33. Black, D. L., Chabot, B. & Steitz, J. A. U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* **42,** 737–750 (1985).

34. Chabot, B., Black, D. L., LeMaster, D. M. & Steitz, J. A. The 3' splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science* **230,** 1344–9 (1985).

35. Black, D. L. & Steitz, J. A. Pre-mRNA splicing in vitro requires intact U4/U6 small nuclear ribonucleoprotein. *Cell* **46,** 697–704 (1986).

36. Berget, S. M. & Robberson, B. L. U1, U2, and U4/U6 small nuclear ribonucleoproteins are required for in vitro splicing but not polyadenylation. *Cell* **46,** 691–696 (1986).

37. Hicks, M. J., Lam, B. J. & Hertel, K. J. Analyzing mechanisms of alternative pre-mRNA splicing using in vitro splicing assays. *Methods* **37,** 306–313 (2005).

38. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136,** 701–718 (2009).

39. Girard, C. *et al.* Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.* **3,** 994 (2012).

40. Saha, K., Fernandez, M. M., Biswas, T., Lumba, C. L. & Ghosh, G. In vitro assembly of an early spliceosome defining both splice sites. *bioRxiv* 292458 (2018). doi:10.1101/292458

41. Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A. & Steitz, J. A. The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site in vitro. *Cell* **33,** 509–518 (1983).

42. Berglund, J. A., Abovich, N. & Rosbash, M. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.* **12,** 858–67 (1998).

43. Wu, S., Romfo, C. M., Nilsen, T. W. & Green, M. R. Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. *Nature* **402,** 832–835 (1999).

44. Guth, S. & Valcárcel, J. Kinetic role for mammalian SF1/BBP in spliceosome assembly and function after polypyrimidine tract recognition by U2AF. *J. Biol. Chem.* **275,** 38059–66 (2000).

45. Graveley, B. R., Hertel, K. J. & Maniatis, T. The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* **7,** 806–18 (2001).

46. Cho, S. *et al.* Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 8233–8 (2011).

47. Sutandy, F. X. R. *et al.* In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. *Genome Res.* (2018). doi:10.1101/gr.229757.117

48. Behzadnia, N. *et al.* Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *EMBO J.* **26,** 1737–48 (2007).

49. Sharma, S., Wongpalee, S. P., Vashisht, A., Wohlschlegel, J. A. & Black, D. L. Stem-loop 4 of U1 snRNA is essential for splicing and interacts with the U2 snRNP-specific SF3A1 protein during spliceosome assembly. *Genes Dev.* **28,** 2518–31 (2014).

50. Konforti, B. B. & Konarska, M. M. U4/U5/U6 snRNP recognizes the 5’ splice site in the absence of U2 snRNP. *Genes Dev.* **8,** 1962–73 (1994).

51. Deckert, J. *et al.* Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol. Cell. Biol.* **26,** 5528–43 (2006).

52. Schmidt, C. *et al.* Mass spectrometry-based relative quantification of proteins in precatalytic and catalytically active spliceosomes by metabolic labeling (SILAC), chemical labeling (iTRAQ), and label-free spectral count. *RNA* **20,** 406–20 (2014).

53. Plaschka, C., Lin, P.-C. & Nagai, K. Structure of a pre-catalytic spliceosome. *Nature* **546,** 617 (2017).

54. Ajuh, P. *et al.* Functional analysis of the human CDC5L complex and identification of its components by mass spectrometry. *EMBO J.* **19,** 6569–81 (2000).

55. Chan, S.-P. & Cheng, S.-C. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J. Biol. Chem.* **280,** 31190–9 (2005).

56. Bessonov, S., Anokhina, M., Will, C. L., Urlaub, H. & Lührmann, R. Isolation of an active step I spliceosome and composition of its RNP core. *Nature* **452,** 846–850 (2008).

57. Bessonov, S. *et al.* Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* **16,** 2384–403 (2010).

58. Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. Structure of a human catalytic step I spliceosome. *Science* **359,** 537–545 (2018).

59. Ruskin, B., Krainer, A. R., Maniatis, T. & Green, M. R. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* **38,** 317–31 (1984).

60. Grabowski, P. J., Padgett, R. A. & Sharp, P. A. Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. *Cell* **37,** 415–27 (1984).

61. Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F. & Sharp, P. A. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* **225,** 898–903 (1984).

62. Ilagan, J. O., Chalkley, R. J., Burlingame, A. L. & Jurica, M. S. Rearrangements within human spliceosomes captured after exon ligation. *RNA* **19,** 400–12 (2013).

63. Tange, T. Ø., Nott, A. & Moore, M. J. The ever-increasing complexities of the exon junction complex. *Curr. Opin. Cell Biol.* **16,** 279–284 (2004).

64. Sträßer, K. *et al.* TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417,** 304–308 (2002).

65. Valadkhan, S. & Jaladat, Y. The spliceosomal proteome: at the heart of the largest cellular ribonucleoprotein machine. *Proteomics* **10,** 4128–41 (2010).

66. Glisovic, T., Bachorik, J. L., Yong, J. & Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* **582,** 1977–1986 (2008).

67. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149,** 1393–1406 (2012).

68. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Hum. Genet.* **135,** 851–867 (2016).

69. Han, S. P., Tang, Y. H. & Smith, R. Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.* **430,** 379–92 (2010).

70. Beyer, A. L., Christensen, M. E., Walker, B. W. & LeStourgeon, W. M. Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell* **11,** 127–138 (1977).

71. van Eekelen, C. A., Riemen, T. & van Venrooij, W. J. Specificity in the interaction of hnRNA and mRNA with proteins as revealed by in vivo cross linking. *FEBS Lett.* **130,** 223–6 (1981).

72. Dreyfuss, G., Choi, Y. D. & Adam, S. A. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Mol. Cell. Biol.* **4,** 1104–14 (1984).

73. Piñol-Roma, S., Choi, Y. D., Matunis, M. J. & Dreyfuss, G. Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev.* **2,** 215–27 (1988).

74. Dreyfuss, G., Kim, V. N. & Kataoka, N. Messenger-RNA-binding proteins and the messages they carry. *Nat. Rev. Mol. Cell Biol.* **3,** 195–205 (2002).

75. Martinez-Contreras, R. *et al.* hnRNP proteins and splicing control. *Adv. Exp. Med. Biol.* **623,** 123–47 (2007).

76. Llorian, M. *et al.* Position-dependent alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat. Struct. Mol. Biol.* **17,** 1114–23 (2010).

77. Huelga, S. C. *et al.* Integrative Genome-wide Analysis Reveals Cooperative Regulation of Alternative Splicing by hnRNP Proteins. *Cell Rep.* **1,** 167–178 (2012).

78. Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W. & Keene, J. D. RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 2495–9 (1991).

79. Zahler, A. M., Lane, W. S., Stolk, J. A. & Roth, M. B. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6,** 837–47 (1992).

80. Roth, M. B., Murphy, C. & Gall, J. G. A monoclonal antibody that recognizes a phosphorylated epitope stains lampbrush chromosome loops and small granules in the amphibian germinal vesicle. *J. Cell Biol.* **111,** 2217–23 (1990).

81. Änkö, M.-L. Regulation of gene expression programmes by serine–arginine rich splicing factors. *Semin. Cell Dev. Biol.* **32,** 11–21 (2014).

82. Han, J. *et al.* SR proteins induce alternative exon skipping through their activities on the flanking constitutive exons. *Mol. Cell. Biol.* **31,** 793–802 (2011).

83. Pandit, S. *et al.* Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol. Cell* **50,** 223–35 (2013).

84. Mayeda, A., Screaton, G. R., Chandler, S. D., Fu, X. D. & Krainer, A. R. Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. *Mol. Cell. Biol.* **19,** 1853–63 (1999).

85. Cáceres, J. F., Misteli, T., Screaton, G. R., Spector, D. L. & Krainer, A. R. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* **138,** 225–38 (1997).

86. Zhu, J. & Krainer, A. R. Pre-mRNA splicing in the absence of an SR protein RS domain. *Genes Dev.* **14,** 3166–78 (2000).

87. Saha, K. *et al.* Exonic unpaired elements modulate pre-mRNA structure for splice site recognition. *bioRxiv* 292003 (2018). doi:10.1101/292003

88. Jankowsky, E. RNA helicases at work: binding and rearranging. *Trends Biochem. Sci.* **36,** 19–29 (2011).

89. Cordin, O., Hahn, D. & Beggs, J. D. Structure, function and regulation of spliceosomal RNA helicases. *Curr. Opin. Cell Biol.* **24,** 431–438 (2012).

90. Cordin, O. & Beggs, J. D. RNA helicases in splicing. *RNA Biol.* **10,** 83–95 (2013).

91. Zhou, Z., Licklider, L. J., Gygi, S. P. & Reed, R. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419,** 182–185 (2002).

92. Rappsilber, J., Ryder, U., Lamond, A. I. & Mann, M. Large-scale proteomic analysis of the human spliceosome. *Genome Res.* **12,** 1231–45 (2002).

93. Chen, Y.-I. G. *et al.* Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. *Nucleic Acids Res.* **35,** 3928–44 (2007).

94. Hegele, A. *et al.* Dynamic Protein-Protein Interaction Wiring of the Human Spliceosome. *Mol. Cell* **45,** 567–580 (2012).

95. Cvitkovic, I. & Jurica, M. S. Spliceosome database: a tool for tracking components of the spliceosome. *Nucleic Acids Res.* **41,** D132-41 (2013).

96. Boesler, C. *et al.* A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat. Commun.* **7,** 11997 (2016).

97. Ulrich, A. K. C., Seeger, M., Schütze, T., Bartlick, N. & Wahl, M. C. Scaffolding in the Spliceosome via Single α Helices. *Structure* **24,** 1972–1983 (2016).

98. Bertram, K. *et al.* Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542,** 318–323 (2017).

99. Zhang, X. *et al.* An Atomic Structure of the Human Spliceosome. *Cell* **169,** 918–929.e14 (2017).

100. Costa, M., Walbott, H., Monachello, D., Westhof, E. & Michel, F. Crystal Structures of a Group II Intron Lariat and Implications for the Spliceosome. *Biophys. J.* **114,** 366a (2018).

101. Zhang, X. *et al.* Structure of the human activated spliceosome in three conformational states. *Cell Res.* **28,** 307–322 (2018).

102. DeHaven, A. C., Norden, I. S. & Hoskins, A. A. Lights, camera, action! Capturing the spliceosome and pre-mRNA splicing with single-molecule fluorescence microscopy. *Wiley Interdiscip. Rev. RNA* **7,** 683–701 (2016).

103. Agafonov, D. E. *et al.* Semiquantitative Proteomic Analysis of the Human Spliceosome via a Novel Two-Dimensional Gel Electrophoresis Method. *Mol. Cell. Biol.* **31,** 2667–2682 (2011).

104. De Conti, L., Baralle, M. & Buratti, E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* **4,** 49–60 (2013).

105. House, A. E. & Lynch, K. W. Regulation of alternative splicing: more than just the ABCs. *J. Biol. Chem.* **283,** 1217–21 (2008).

106. Fu, X.-D., Ares, M. & Jr. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* **15,** 689–701 (2014).

107. Tilgner, H. *et al.* Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* **16,** 996–1001 (2009).

108. Schwartz, S., Meshorer, E. & Ast, G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **16,** 990–995 (2009).

109. Gueroussov, S. *et al.* Regulatory Expansion in Mammals of Multivalent hnRNP Assemblies that Globally Control Alternative Splicing. *Cell* **170,** 324–339.e23 (2017).

110. Sammeth, M., Foissac, S. & Guigó, R. A general definition and nomenclature for alternative splicing events. *PLoS Comput. Biol.* **4,** e1000147 (2008).

111. Pohl, M., Bortfeldt, R. H., Grützmann, K. & Schuster, S. Alternative splicing of mutually exclusive exons—A review. *Biosystems* **114,** 31–38 (2013).

112. Shin, C. & Manley, J. L. Cell signalling and the control of pre-mRNA splicing. *Nat. Rev. Mol. Cell Biol.* **5,** 727–738 (2004).

113. Lynch, K. W. Regulation of alternative splicing by signal transduction pathways. *Adv. Exp. Med. Biol.* **623,** 161–74 (2007).

114. Heyd, F. & Lynch, K. W. Degrade, move, regroup: signaling control of splicing proteins. *Trends Biochem. Sci.* **36,** 397–404 (2011).

115. Naro, C. & Sette, C. Phosphorylation-mediated regulation of alternative splicing in cancer.

Int. J. Cell Biol. **2013,** 151839 (2013).

116. Mermoud, J. E., Cohen, P. T. & Lamond, A. I. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO J.* **13,** 5679–5688 (1994).

117. Misteli, T. & Spector, D. L. Protein phosphorylation and the nuclear organization of pre-mRNA splicing. *Trends Cell Biol.* **7,** 135–138 (1997).

118. Stamm, S. Regulation of alternative splicing by reversible protein phosphorylation. *J. Biol. Chem.* **283,** 1223–7 (2008).

119. Misteli, T. *et al.* Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *J. Cell Biol.* **143,** 297–307 (1998).

120. Xiang, S. *et al.* Phosphorylation Drives a Dynamic Switch in Serine/Arginine-Rich Proteins. *Structure* **21,** 2162–2174 (2013).

121. Rie Manceau, V. *et al.* Major phosphorylation of SF1 on adjacent Ser-Pro motifs enhances interaction with U2AF. doi:10.1111/j.1742-4658.2005.05091.x

122. Wang, X. *et al.* Phosphorylation of splicing factor SF1 on Ser20 by cGMP-dependent protein kinase regulates spliceosome assembly. *EMBO J.* **18,** 4549–59 (1999).

123. Gonçalves, V. *et al.* Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells. *RNA* **20,** 474–82 (2014).

124. Aubol, B. E., Keshwani, M. M., Fattet, L. & Adams, J. A. Mobilization of a splicing factor through a nuclear kinase-kinase complex. *Biochem. J.* **475,** 677–690 (2018).

125. Huang, Y., Yario, T. A. & Steitz, J. A. A molecular link between SR protein dephosphorylation and mRNA export. *Proc. Natl. Acad. Sci.* **101,** 9666–9670 (2004).

126. van der Houven van Oordt, W. *et al.* The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J. Cell Biol.* **149,** 307–16 (2000).

127. Allemand, E. *et al.* Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3605–10 (2005).

128. Guil, S., Long, J. C. & Cáceres, J. F. hnRNP A1 relocalization to the stress granules reflects a role in the stress response. *Mol. Cell. Biol.* **26,** 5744–58 (2006).

129. Wang, C. *et al.* Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.* **12,** 1409–1414 (1998).

130. de Graaf, K. *et al.* The protein kinase DYRK1A phosphorylates the splicing factor SF3b1/SAP155 at Thr434, a novel in vivo phosphorylation site. *BMC Biochem.* **7,** 7 (2006).

131. Schneider, M. *et al.* Human PRP4 kinase is required for stable tri-snRNP association during spliceosomal B complex formation. *Nat. Struct. Mol. Biol.* **17,** 216–221 (2010).

132. Gräub, R. *et al.* Cell cycle-dependent phosphorylation of human CDC5 regulates RNA processing. *Cell Cycle* **7,** 1795–803 (2008).

133. Hirose, Y., Tacke, R. & Manley, J. L. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev.* **13,** 1234–9 (1999).

134. Das, R. *et al.* SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol. Cell* **26,** 867–81 (2007).

135. de la Mata, M. & Kornblihtt, A. R. RNA polymerase II C-terminal domain mediates

regulation of alternative splicing by SRp20. *Nat. Struct. Mol. Biol.* **13,** 973–980 (2006).

136. Rajan, P. *et al.* The RNA-binding and adaptor protein Sam68 modulates signal-dependent splicing and transcriptional activity of the androgen receptor. *J. Pathol.* **215,** 67–77 (2008).

137. Zhou, Z. *et al.* The Akt-SRPK-SR Axis Constitutes a Major Pathway in Transducing EGF Signaling to Regulate Alternative Splicing in the Nucleus. *Mol. Cell* **47,** 422–433 (2012).

138. Cargnello, M. & Roux, P. P. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol. Mol. Biol. Rev.* **75,** 50–83 (2011).

139. Weg-Remers, S., Ponta, H., Herrlich, P. & König, H. Regulation of alternative pre-mRNA splicing by the ERK MAP-kinase pathway. *EMBO J.* **20,** 4194–203 (2001).

140. Matter, N., Herrlich, P. & König, H. Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature* **420,** 691–695 (2002).

141. Al-Ayoubi, A. M., Zheng, H., Liu, Y., Bai, T. & Eblen, S. T. Mitogen-activated protein kinase phosphorylation of splicing factor 45 (SPF45) regulates SPF45 alternative splicing site utilization, proliferation, and cell adhesion. *Mol. Cell. Biol.* **32,** 2880–93 (2012).

142. Sinha, R. *et al.* Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. *Mol. Cell. Biol.* **30,** 2762–74 (2010).

143. Wang, F. *et al.* SPSB1-mediated HnRNP A1 ubiquitylation regulates alternative splicing and cell migration in EGF signaling. *Cell Res.* **27,** 540–558 (2017).

144. Webby, C. J. *et al.* Jmjd6 Catalyses Lysyl-Hydroxylation of U2AF65, a Protein Associated with RNA Splicing. *Science (80-.).* **325,** 90–93 (2009).

145. Faustino, N. A. & Cooper, T. A. Pre-mRNA splicing and human disease. *Genes Dev.* **17,** 419–37 (2003).

146. Padgett, R. A. New connections between splicing and human disease. *Trends Genet.* **28,** 147–54 (2012).

147. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **17,** 19–32 (2016).

148. Chabot, B. & Shkreta, L. Defective control of pre-messenger RNA splicing in human disease. *J. Cell Biol.* **212,** 13–27 (2016).

149. Takeshima, Y. *et al.* Mutation spectrum of the dystrophin gene in 442 Duchenne/Becker muscular dystrophy cases from one Japanese referral center. *J. Hum. Genet.* **55,** 379–388 (2010).

150. Fletcher, S. *et al.* Antisense suppression of donor splice site mutations in the dystrophin gene transcript. *Mol. Genet. genomic Med.* **1,** 162–73 (2013).

151. Jung, H. *et al.* Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat. Genet.* **47,** 1242–1248 (2015).

152. Venables, J. P. Aberrant and alternative splicing in cancer. *Cancer Res.* **64,** 7647–54 (2004).

153. David, C. J. & Manley, J. L. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* **24,** 2343–64 (2010).

154. Zhang, J. & Manley, J. L. Misregulation of Pre-mRNA Alternative Splicing in Cancer. *Cancer Discov.* **3,** 1228–1237 (2013).

155. Ladomery, M. Aberrant alternative splicing is another hallmark of cancer. *Int. J. Cell Biol.* **2013,** 463786 (2013).

156. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* **33,** 5311–5318 (2014).

157. Martinez-Montiel, N., Rosas-Murrieta, N., Anaya Ruiz, M., Monjaraz-Guzman, E. & Martinez-Contreras, R. Alternative Splicing as a Target for Cancer Treatment. *Int. J. Mol. Sci.* **19,** 545 (2018).

158. Grosso, A. R., Martins, S. & Carmo-Fonseca, M. The emerging role of splicing factors in cancer. *EMBO Rep.* **9,** 1087–1093 (2008).

159. Dvinge, H., Kim, E., Abdel-Wahab, O. & Bradley, R. K. RNA splicing factors as oncoproteins and tumour suppressors. *Nat. Rev. Cancer* **16,** 413–430 (2016).

160. Anczuków, O. & Krainer, A. R. Splicing-factor alterations in cancers. *RNA* **22,** 1285–301 (2016).

161. Sveen, A., Kilpinen, S., Ruusulehto, A., Lothe, R. A. & Skotheim, R. I. Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene* **35,** 2413–2427 (2016).

162. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–74 (2011).

163. Lee, S. C.-W. & Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat. Med.* **22,** 976–86 (2016).

164. Salton, M. & Misteli, T. Small Molecule Modulators of Pre-mRNA Splicing in Cancer Therapy. *Trends Mol. Med.* **22,** 28–37 (2016).

165. Jyotsana, N. & Heuser, M. Exploiting differential RNA splicing patterns: a potential new group of therapeutic targets in cancer. *Expert Opin. Ther. Targets* **22,** 107–121 (2018).

166. Urbanski, L. M., Leclair, N. & Anczuków, O. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. *Wiley Interdiscip. Rev. RNA* e1476 (2018). doi:10.1002/wrna.1476

167. Chung, F. F.-L. *et al.* Jerantinine A induces tumor-specific cell death through modulation of splicing factor 3b subunit 1 (SF3B1). *Sci. Rep.* **7,** 42504 (2017).

168. Seiler, M. *et al.* H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **24,** 497–504 (2018).

169. Liu, T. *et al.* TRA2A Promoted Paclitaxel Resistance and Tumor Progression in Triple-Negative Breast Cancers via Regulating Alternative Splicing. *Mol. Cancer Ther.* **16,** 1377–1388 (2017).

170. Karni, R. *et al.* The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nat. Struct. Mol. Biol.* **14,** 185–93 (2007).

171. Anczuków, O. *et al.* The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. *Nat. Struct. Mol. Biol.* **19,** 220–228 (2012).

172. Das, S. & Krainer, A. R. Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol. Cancer Res.* **12,** 1195–204 (2014).

173. Adler, A. S. *et al.* An integrative analysis of colon cancer identifies an essential function for PRPF6 in tumor growth. *Genes Dev.* **28,** 1068–84 (2014).

174. David, C. J., Chen, M., Assanah, M., Canoll, P. & Manley, J. L. HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature* **463,** 364–368 (2010).

175. Golan-Gerstl, R. *et al.* Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. *Cancer Res.* **71,** 4464–72 (2011).

176. Dai, P. *et al.* Unraveling Molecular Differences of Gastric Cancer by Label-Free Quantitative Proteomics Analysis. *Int. J. Mol. Sci.* **17,** (2016).

177. Fei, T. *et al.* Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc. Natl. Acad. Sci. U. S. A.* **114,** E5207–E5215 (2017).

178. Zhou, X. *et al.* HnRNP-L promotes prostate cancer progression by enhancing cell cycling and inhibiting apoptosis. *Oncotarget* **8,** 19342–19353 (2017).

179. Datar, K. V, Dreyfuss, G. & Swanson, M. S. The human hnRNP M proteins: identification of a methionine/arginine-rich repeat motif in ribonucleoproteins. *Nucleic Acids Res.* **21,** 439–46 (1993).

180. Aidinis, V., Sekeris, C. E. & Guialis, A. Two immunologically related polypeptides of 72/74 kDa specify a novel 70-100S heterogeneous nuclear RNP. *Nucleic Acids Res.* **23,** 2742–53 (1995).

181. Kafasla, P., Patrinou-Georgoula, M. & Guialis, A. The 72/74-kDa polypeptides of the 70-110 S large heterogeneous nuclear ribonucleoprotein complex (LH-nRNP) represent a discrete subset of the hnRNP M protein family. *Biochem. J.* **350 Pt 2,** 495–503 (2000).

182. Kafasla, P., Patrinou-Georgoula, M., Lewis, J. D. & Guialis, A. Association of the 72/74-kDa proteins, members of the heterogeneous nuclear ribonucleoprotein M group, with the pre-mRNA at early stages of spliceosome assembly. *Biochem. J.* **363,** 793–9 (2002).

183. Gattoni, R. *et al.* The human hnRNP-M proteins: structure and relation with early heat shock-induced splicing arrest and chromosome mapping. *Nucleic Acids Res.* **24,** 2535–42 (1996).

184. Hase, M. E., Yalamanchili, P. & Visa, N. The Drosophila heterogeneous nuclear ribonucleoprotein M protein, HRP59, regulates alternative splicing and controls the production of its own mRNA. *J. Biol. Chem.* **281,** 39135–41 (2006).

185. Hovhannisyan, R. H. & Carstens, R. P. Heterogeneous ribonucleoprotein m is a splicing regulatory protein that can enhance or silence splicing of alternatively spliced exons. *J. Biol. Chem.* **282,** 36265–74 (2007).

186. Xu, Y. *et al.* Cell type-restricted activity of hnRNPM promotes breast cancer metastasis via regulating alternative splicing. *Genes Dev.* **28,** 1191–203 (2014).

187. Passacantilli, I., Frisone, P., De Paola, E., Fidaleo, M. & Paronetto, M. P. hnRNPM guides an alternative splicing program in response to inhibition of the PI3K/AKT/mTOR pathway in Ewing sarcoma cells. *Nucleic Acids Res.* **45,** 12270–12284 (2017).

188. Llères, D., Denegri, M., Biggiogera, M., Ajuh, P. & Lamond, A. I. Direct interaction between hnRNP-M and CDC5L/PLRG1 proteins affects alternative splice site choice. *EMBO Rep.* **11,** 445–451 (2010).

189. Sun, H. *et al.* HnRNPM and CD44s expression affects tumor aggressiveness and predicts poor prognosis in breast cancer with axillary lymph node metastases. *Genes, Chromosom. Cancer* **56,** 598–607 (2017).

190. Bajenova, O. V *et al.* Heterogeneous RNA-binding protein M4 is a receptor for carcinoembryonic antigen in Kupffer cells. *J. Biol. Chem.* **276,** 31067–73 (2001).

191. Thomas, P., Forse, R. A. & Bajenova, O. Carcinoembryonic antigen (CEA) and its receptor hnRNP M are mediators of metastasis and the inflammatory response in the liver. *Clin. Exp.*

Metastasis **28,** 923–932 (2011).

192. Bajenova, O. *et al.* Carcinoembryonic antigen promotes colorectal cancer progression by targeting adherens junction complexes. *Exp. Cell Res.* **324,** 115–123 (2014).

193. Chen, S. *et al.* Identification of HnRNP M as a novel biomarker for colorectal carcinoma by quantitative proteomics. *Am. J. Physiol. Liver Physiol.* **306,** G394–G403 (2014).

194. Bjersand, K., Seidal, T., Sundström-Poromaa, I., Åkerud, H. & Skirnisdottir, I. The clinical and prognostic correlation of HRNPM and SLC1A5 in pathogenesis and prognosis in epithelial ovarian cancer. *PLoS One* **12,** e0179363 (2017).

195. Grossman, R. L. *et al.* Toward a Shared Vision for Cancer Genomic Data. *N. Engl. J. Med.* **375,** 1109–1112 (2016).

196. Uhlén, M. *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4,** 1920–32 (2005).

197. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28,** 1248–1250 (2010).

198. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357,** eaan2507 (2017).

199. Győrffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123,** 725–731 (2010).

200. Győrffy, B., Surowiak, P., Budczies, J. & Lánczky, A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One* **8,** e82241 (2013).

201. Szász, A. M. *et al.* Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. *Oncotarget* **7,** 49322–49333 (2016).

202. Weissbach, L. *et al.* Identification of a human rasGAP-related protein containing calmodulin-binding motifs. *J. Biol. Chem.* **269,** 20517–21 (1994).

203. Hart, M. J., Callow, M. G., Souza, B. & Polakis, P. IQGAP1, a calmodulin-binding protein with a rasGAP-related domain, is a potential effector for cdc42Hs. *EMBO J.* **15,** 2997–3005 (1996).

204. Yamashiro, S., Noguchi, T. & Mabuchi, I. Localization of Two IQGAPs in Cultured Cells and Early Embryos of Xenopus laevis. **50,** 36–50 (2003).

205. Chew, C. S., Okamoto, C. T., Chen, X. & Qin, H. Y. IQGAPs are differentially expressed and regulated in polarized gastric epithelial cells. *Am. J. Physiol. Liver Physiol.* **288,** G376–G387 (2005).

206. White, C. D., Brown, M. D. & Sacks, D. B. IQGAPs in cancer: A family of scaffold proteins underlying tumorigenesis. *FEBS Lett.* **583,** 1817–1824 (2009).

207. Smith, J. M., Hedman, A. C. & Sacks, D. B. IQGAPs choreograph cellular signaling from the membrane to the nucleus. *Trends Cell Biol.* **25,** 171–184 (2015).

208. Hedman, A. C., Smith, J. M. & Sacks, D. B. The biology of IQGAP proteins: beyond the cytoskeleton. *EMBO Rep.* **16,** 427–446 (2015).

209. Briggs, M. W., Li, Z. & Sacks, D. B. IQGAP1-mediated stimulation of transcriptional co-activation by β-catenin is modulated by calmodulin. *J. Biol. Chem.* **277,** 7453–7465 (2002).

210. Fukata, M. *et al.* Rac1 and Cdc42 Capture Microtubules through IQGAP1 and CLIP-170. *Cell* **109,** 873–885 (2002).

211. Watanabe, T. *et al.* Interaction with IQGAP1 links APC to Rac1, Cdc42, and actin filaments during cell polarization and migration. *Dev. Cell* **7,** 871–883 (2004).

212. Fukata, M. *et al.* Regulation of cross-linking of actin filament by IQGAP1, a target for Cdc42. *J. Biol. Chem.* **272,** 29579–83 (1997).

213. Noritake, J. IQGAP1: a key regulator of adhesion and migration. *J. Cell Sci.* **118,** 2085–2092 (2005).

214. Mataraza, J. M. *et al.* IQGAP1 promotes cell motility and invasion. *J. Biol. Chem.* **278,** 41237–45 (2003).

215. Watanabe, T. *et al.* Interaction with IQGAP1 Links APC to Rac1, Cdc42, and Actin Filaments during Cell Polarization and Migration. *Dev. Cell* **7,** 871–883 (2004).

216. Mataraza, J. M., Li, Z., Jeong, H.-W., Brown, M. D. & Sacks, D. B. Multiple proteins mediate IQGAP1-stimulated cell migration. *Cell. Signal.* **19,** 1857–1865 (2007).

217. White, C. D., Erdemir, H. H. & Sacks, D. B. IQGAP1 and its binding proteins control diverse biological functions. *Cell. Signal.* **24,** 826–834 (2012).

218. Good, M. C., Zalatan, J. G. & Lim, W. A. Scaffold proteins: Hubs for controlling the flow of cellular information. *Science (80-.).* **332,** 680–686 (2011).

219. Pan, C. Q., Sudol, M., Sheetz, M. & Low, B. C. Modularity and functional plasticity of scaffold proteins as p(l)acemakers in cell signaling. *Cell. Signal.* **24,** 2143–2165 (2012).

220. Langeberg, L. K. & Scott, J. D. Signalling scaffolds and local organization of cellular behaviour. *Nat. Rev. Mol. Cell Biol.* **16,** 232–244 (2015).

221. Matsunaga, H., Kubota, K., Inoue, T., Isono, F. & Ando, O. IQGAP1 selectively interacts with K-Ras but not with H-Ras and modulates K-Ras function. *Biochem. Biophys. Res. Commun.* **444,** 360–364 (2014).

222. Ren, J.-G., Li, Z. & Sacks, D. B. IQGAP1 modulates activation of B-Raf. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 10465–9 (2007).

223. Ren, J.-G., Li, Z. & Sacks, D. B. IQGAP1 integrates Ca2+/calmodulin and B-Raf signaling. *J. Biol. Chem.* **283,** 22972–82 (2008).

224. Roy, M., Li, Z. & Sacks, D. B. IQGAP1 is a scaffold for mitogen-activated protein kinase signaling. *Mol. Cell. Biol.* **25,** 7940–52 (2005).

225. Roy, M., Li, Z. & Sacks, D. B. IQGAP1 binds ERK2 and modulates its activity. *J. Biol. Chem.* **279,** 17329–37 (2004).

226. Bardwell, A. J., Lagunes, L., Zebarjedi, R. & Bardwell, L. The WW domain of the scaffolding protein IQGAP1 is neither necessary nor sufficient for binding to the MAPKs ERK1 and ERK2. *J. Biol. Chem.* **292,** 8750–8761 (2017).

227. Vetterkind, S., Lin, Q. Q. & Morgan, K. G. A novel mechanism of ERK1/2 regulation in smooth muscle involving acetylation of the ERK1/2 scaffold IQGAP1. *Sci. Rep.* **7,** 9302 (2017).

228. Goto, T. *et al.* IQGAP1 protein regulates nuclear localization of β-Catenin via importin-β5 protein in wnt signaling. *J. Biol. Chem.* **288,** 36351–36360 (2013).

229. Goto, T. *et al.* IQGAP1 Functions as a Modulator of Dishevelled Nuclear Localization in

Wnt Signaling. *PLoS One* **8,** 1–11 (2013).

230. Johnson, M., Sharma, M., Brocardo, M. G. & Henderson, B. R. IQGAP1 translocates to the nucleus in early S-phase and contributes to cell cycle progression after DNA replication arrest. *Int. J. Biochem. Cell Biol.* **43,** 65–73 (2011).

231. Bielak-Zmijewska, A., Kolano, A., Szczepanska, K., Maleszewski, M. & Borsuk, E. Cdc42 protein acts upstream of IQGAP1 and regulates cytokinesis in mouse oocytes and embryos. *Dev. Biol.* **322,** 21–32 (2008).

232. Kuroda, S. *et al.* Role of IQGAP1, a target of the small GTPases Cdc42 and Rac1, in regulation of E-cadherin- mediated cell-cell adhesion. *Science* **281,** 832–5 (1998).

233. Takaine, M., Numata, O. & Nakano, K. Fission yeast IQGAP arranges actin filaments into the cytokinetic contractile ring. *EMBO J.* **28,** 3117–31 (2009).

234. Rigothier, C. *et al.* Nuclear translocation of IQGAP1 protein upon exposure to puromycin aminonucleoside in cultured human podocytes: ERK pathway involvement. *Cell. Signal.* **28,** 1470–1478 (2016).

235. Erdemir, H. H., Li, Z. & Sacks, D. B. IQGAP1 binds to estrogen receptor-α and modulates its function. *J. Biol. Chem.* **289,** 9100–12 (2014).

236. Kim, J.-H. *et al.* Identification and Functional Studies of a New Nrf2 Partner IQGAP1: A Critical Role in the Stability and Transactivation of Nrf2. *Antioxid. Redox Signal.* **19,** 89–101 (2013).

237. Sharma, S. *et al.* Dephosphorylation of the nuclear factor of activated T cells (NFAT) transcription factor is regulated by an RNA-protein scaffold complex. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 11381–6 (2011).

238. Villacé, P., Marión, R. M. & Ortín, J. The composition of Staufen-containing RNA granules from human cells indicates their role in the regulated transport and translation of messenger RNAs. *Nucleic Acids Res.* **32,** 2411–2420 (2004).

239. Meyer, E. L. & Gavis, E. R. Staufen does double duty. *Nat. Struct. Mol. Biol.* **12,** 291–292 (2005).

240. Kim, Y. K. *et al.* Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J.* **26,** 2670–81 (2007).

241. Gehring, N. H., Lamprinaki, S., Kulozik, A. E. & Hentze, M. W. Disassembly of exon junction complexes by PYM. *Cell* **137,** 536–48 (2009).

242. Popp, M. W.-L. & Maquat, L. E. Organizing Principles of Mammalian Nonsense-Mediated mRNA Decay. *Annu. Rev. Genet.* **47,** 139–165 (2013).

243. Takeda, S. *et al.* Role of a tyrosine phosphorylation of SMG-9 in binding of SMG-9 to IQGAP and the NMD complex. *Biochem. Biophys. Res. Commun.* **410,** 29–33 (2011).

244. Johnson, M., Sharma, M. & Henderson, B. R. IQGAP1 regulation and roles in cancer. *Cell. Signal.* **21,** 1471–1478 (2009).

245. Sanchez-Laorden, B., Viros, A. & Marais, R. Mind the IQGAP. *Cancer Cell* **23,** 715–717 (2013).

246. Jadeski, L., Mataraza, J. M., Jeong, H.-W., Li, Z. & Sacks, D. B. IQGAP1 stimulates proliferation and enhances tumorigenesis of human breast epithelial cells. *J. Biol. Chem.* **283,** 1008–17 (2008).

247. Sun, W. *et al.* Identification of differentially expressed genes in human lung squamous cell

carcinoma using suppression subtractive hybridization. *Cancer Lett.* **212,** 83–93 (2004).

248. Bertucci, F. *et al.* Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. *Oncogene* **23,** 1377–1391 (2004).

249. Ouyang, X. *et al.* Activator protein-1 transcription factors are associated with progression and recurrence of prostate cancer. *Cancer Res.* **68,** 2132–44 (2008).

250. Sugimoto, N. *et al.* IQGAP1 , a negative regulator of cell-cell adhesion, is upregulated by gene amplification at 15q26 in gastric cancer cell lines HSC39 and 40A. *J. Hum. Genet.* **46,** 21–25 (2001).

251. Dong, P.-X. *et al.* Silencing of IQGAP1 by shRNA inhibits the invasion of ovarian carcinoma HO-8910PM cells in vitro. *J. Exp. Clin. Cancer Res.* **27,** 77 (2008).

252. Hu, B. *et al.* ADP-ribosylation factor 6 regulates glioma cell invasion through the IQ-domain GTPase-activating protein 1-Rac1-mediated pathway. *Cancer Res.* **69,** 794–801 (2009).

253. Liang, Z. *et al.* SUMOylation of IQGAP1 promotes the development of colorectal cancer. *Cancer Lett.* **411,** 90–99 (2017).

254. Li, S., Wang, Q., Chakladar, A., Bronson, R. T. & Bernards, A. Gastric hyperplasia in mice lacking the putative Cdc42 effector IQGAP1. *Mol. Cell. Biol.* **20,** 697–701 (2000).

255. Takemoto, H. *et al.* Localization of IQGAP1 is inversely correlated with intercellular adhesion mediated by e-cadherin in gastric cancers. *Int. J. Cancer* **91,** 783–788 (2001).

256. Smoot, D. T. *et al.* Human gastric epithelial cell lines derived from primary cultures of normal gastric epithelial cells. *Gastroenterology* **118,** A540–A541 (2000).

257. Akiyama, S. *et al.* Characteristics of three human gastric cancer cell lines, NU-GC-2, NU-GC-3 and NU-GC-4. *Jpn. J. Surg.* **18,** 438–446 (1988).

258. Nakashio, T. *et al.* The association of metastasis with the expression of adhesion molecules in cell lines derived from human gastric cancer. *Anticancer Res.* **17,** 293–9 (1997).

259. Nakashio, T. *et al.* Adhesion molecules and TGF-β1 are involved in the peritoneal dissemination of NUGC-4 human gastric cancer cells. *Int. J. Cancer* **70,** 612–618 (1997).

260. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72,** 248–254 (1976).

261. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6,** 359–362 (2009).

262. Andrew Thompson, † *et al.* Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. (2003). doi:10.1021/AC0262560

263. Breitwieser, F. P. *et al.* General Statistical Modeling of Data from Protein Relative Expression Isobaric Tags. *J. Proteome Res.* **10,** 2758–2766 (2011).

264. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3,** 1–25 (2004).

265. Hughes, C. S. *et al.* Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10,** 757 (2014).

266. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–2308 (2013).

267. Trevino, A. E. & Zhang, F. Genome Editing Using Cas9 Nickases. *Methods Enzymol.* **546,**

161–174 (2014).

268. Zhang Lab. Optimized CRISPR Design. (2013). Available at: http://crispr.mit.edu/.

269. Chiang, T.-W. W., le Sage, C., Larrieu, D., Demir, M. & Jackson, S. P. CRISPR-Cas9D10A nickase-based genotypic and phenotypic screening to enhance genome editing. *Sci. Rep.* **6,** 24356 (2016).

270. Crowley, L. C., Christensen, M. E. & Waterhouse, N. J. Measuring Survival of Adherent Cells with the Colony-Forming Assay. *Cold Spring Harb. Protoc.* **2016,** pdb.prot087171 (2016).

271. Krishan, A. Rapid flow cytofluorometric analysis of mammalian cell cycle by propidium iodide staining. *J. Cell Biol.* **66,** 188–93 (1975).

272. Riccardi, C. & Nicoletti, I. Analysis of apoptosis by propidium iodide staining and flow cytometry. *Nat. Protoc.* **1,** 1458–1461 (2006).

273. Damianov, A. *et al.* Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR. *Cell* **165,** 606–19 (2016).

274. Dominski, Z. & Kole, R. Selection of splice sites in pre-mRNAs with short internal exons. *Mol. Cell. Biol.* **11,** 6075–83 (1991).

275. Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J. M. & Pascual-Montano, A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol.* **8,** R3 (2007).

276. Nogales-Cadenas, R. *et al.* GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res.* **37,** W317–W322 (2009).

277. Tabas-Madrid, D., Nogales-Cadenas, R. & Pascual-Montano, A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res.* **40,** W478–W483 (2012).

278. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25,** 25–29 (2000).

279. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45,** D331–D338 (2017).

280. Cvitkovic, I. & Jurica, M. S. Spliceosome Database: a tool for tracking components of the spliceosome. *Nucleic Acids Res.* **41,** D132–D141 (2013).

281. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46,** D649–D655 (2018).

282. Gao, X. *et al.* FgPrp4 Kinase Is Important for Spliceosome B-Complex Activation and Splicing Efficiency in Fusarium graminearum. *PLOS Genet.* **12,** e1005973 (2016).

283. Bardoni, B. *et al.* 82-FIP, a novel FMRP (Fragile X Mental Retardation Protein) interacting protein, shows a cell cycle-dependent intracellular localization. *Hum. Mol. Genet.* **12,** 1689–1698 (2003).

284. Bardoni, B., Schenck, A. & Louis Mandel, J. A Novel RNA-binding Nuclear Protein That Interacts With the Fragile X Mental Retardation (FMR1) Protein. *Hum. Mol. Genet.* **8,** 2557–2566 (1999).

285. Bardoni, B. *et al.* NUFIP1 (nuclear FMRP interacting protein 1) is a nucleocytoplasmic shuttling protein associated with active synaptoneurosomes. *Exp. Cell Res.* **289,** 95–107

(2003).

286. Lahiri, S. & Futerman, A. H. LASS5 is a bona fide dihydroceramide synthase that selectively utilizes palmitoyl-CoA as acyl donor. *J. Biol. Chem.* **280,** 33735–8 (2005).

287. Senkal, C. E., Ponnusamy, S., Bielawski, J., Hannun, Y. A. & Ogretmen, B. Antiapoptotic roles of ceramide-synthase-6-generated C $_{16}$ -ceramide *via* selective regulation of the ATF6/CHOP arm of ER-stress-response pathways. *FASEB J.* **24,** 296–308 (2010).

288. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–2308 (2013).

289. Menyhárt, O. *et al.* Guidelines for the selection of functional assays to evaluate the hallmarks of cancer. *Biochim. Biophys. Acta - Rev. Cancer* **1866,** 300–319 (2016).

290. Johnson, M., Sharma, M., Brocardo, M. G. & Henderson, B. R. IQGAP1 translocates to the nucleus in early S-phase and contributes to cell cycle progression after DNA replication arrest. *Int. J. Biochem. Cell Biol.* **43,** 65–73 (2011).

291. Villace, P., Marión, R. M. & Ortín, J. The composition of Staufen-containing RNA granules from human cells indicates their role in the regulated transport and translation of messenger RNAs. *Nucleic Acids Res.* **32,** 2411–2420 (2004).

292. Wang, J.-B., Sonn, R., Tekletsadik, Y. K., Samorodnitsky, D. & Osman, M. A. IQGAP1 regulates cell proliferation through a novel CDC42-mTOR pathway. *J. Cell Sci.* **122,** 2024–2033 (2009).