Catarina Rodrigues Nunes

A Catalogue of Stress Granules' Components: Implications for Neurodegeneration

UNIVERSIDADE DO ALGARVE

Departamento de Ciências Biomédicas e Medicina

2019

Catarina Rodrigues Nunes

A Catalogue of Stress Granules' Components: Implications for Neurodegeneration

Master in Oncobiology – Molecular Mechanisms of Cancer

This work was done under the supervision of:

Clévio Nóbrega, Ph.D

UNIVERSIDADE DO ALGARVE

Departamento de Ciências Biomédicas e Medicina

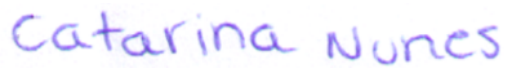
2019

A catalogue of Stress Granules' Components: Implications for neurodegeneration

Declaração de autoria de trabalho

Declaro ser a autora deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam na listagem de referências incluída.

I declare that I am the author of this work, that is original and unpublished. Authors and works consulted are properly cited in the text and included in the list of references.



(Catarina Nunes)

Part of the results of this thesis were published in Nunes,C.; Mestre,I.; Marcelo,A. *et al*. MSGP: the first database of the protein components of the mammalian stress granules. *Database* (2019) Vol. 2019. (In annex A).

ACKNOWLEDGEMENTS

A realização desta tese marca o final de uma etapa académica muito especial e que jamais irei esquecer. No entanto, a finalização desta etapa teria sido impossível sem o apoio de certas pessoas, às quais aqui deixo o meu mais sincero agradecimento.

Ao meu orientador, Professor Doutor Clévio Nóbrega, por ter sido o meu mentor durante este período, pela maneira como me acolheu no seu grupo e pela excelência científica e profissionalismo que sempre demonstrou. Por toda a generosidade, motivação, apoio e infinita paciência, o meu mais profundo agradecimento.

Deixo também um agradecimento muito especial à Doutora Isa Mestre, por todo o apoio, disponibilidade, amabilidade e paciência.

Ás minhas colegas, Adriana Marcelo, Marta Santos e Rebekah Koppenol, um muito obrigada pelo acolhecimento, apoio e amizade. Seria impossível ter pedido um melhor ambiente para trabalhar. Desejo-vos os maiores sucessos.

Aos meus colegas do Mestrado em Oncobiologia, um obrigada por toda a aprendizagem e troca de conhecimentos. Um obrigada especial a quem foi a minha amiga de todas as horas, Marina Baessa, por todo o apoio e amizade. Estendo também este agradecimento a todos os professores deste mestrado, pelo apoio e constante motivação.

Um grande agradecimento à minha família e amigos, que sempre estiveram presentes, tanto neste período como sempre.

ABSTRACT

Stress granules (SGs) are irregularly shaped foci constituted by a variety of different types of RNAs, proteins, factors involved in translation and signaling molecules. These granules form transiently in response to a variety of stress stimuli, such as viral infections or translation blocking drugs, and facilitate stress response by sequestering mRNAs and proteins and affecting the translation of RNA transcripts.

SGs are necessary for stress response, however, abnormalities in SGs functioning and the association of their formation with the aggregation of many pathological proteins, have been linked to pathological changes in several human diseases, such as neurodegenerative diseases, cancer, and aging.

The aim of this work was to identify all the proteins described as SGs components in mammalian cells and to build an online open access database with the all information collected.

The SGs components were identified through an exhaustive study search (PubMed) and further details about each component were retrieved using public databases. Moreover, a transcriptomic analysis was performed analyzing the gene expression data for each SGs component in different neurodegenerative diseases.

We identified 464 proteins as components of mammalian cells SGs, from which 253 were classified as RNA-binding proteins (RBPs). Details about each protein, such as molecular function, link to disease, cell type where they were detected and the stress stimuli used to induce SGs assembly, were also collected and are available in the database. Through a transcriptomic analysis, where “disease vs control” groups were compared and analyzed, a vast majority of SGs proteins were found to be differentially expressed in different neurodegenerative diseases.

All the information collected was used to build the Mammalian Stress Granules Proteome (MSGP), available at https://msgp.pt/, an important tool for researchers in this area of growing interest, being the first database to list and provide information about all SGs proteins identified so far.

Keywords: Stress granules; aggregation; components, proteins, mammalian; neurodegeneration; gene profiling; database

RESUMO

Em condições de stresse, os organismos eucariotas possuem um conjunto de mecanismos para o combater e assim restabelecer a homeostasia celular. Nesta resposta inclui-se a formação dos grânulos de stresse, que são estruturas de formato irregular, sem membrana e formados de forma transiente, que sequestram mRNAs e proteínas, afectando a tradução, e assim facilitando a resposta celular ao stresse.

Os grânulos de stresse são constituidos por uma grande variedade de componentes: proteínas, mRNAs e microRNAs, componentes da tradução e moléculas de sinalização. No entanto, o perfil proteíco destes grânulos varia consoante o tipo celular onde se formam e o tipo de stresse celular, como por exemplo, a privação nutricional, infecções virais, radiações UV, entre outros.

Para a formação dos grânulos de stresse são necessárias proteínas ligantes a RNA (*RNA-binding proteins – RBPs*). As RBPs são proteínas envolvidas numa variedade de processos relacionados com RNA, como o *splicing* alternativo e a biogénese de miRNA, possuindo um papel muito importante na função neuronal. Estas proteínas, para além de fazerem parte da constituição dos grânulos de stresse, são também fundamentais para a sua formação.

Tal como já foi referido a formação dos grânulos de stresse é temporária, uma vez que quando o stresse desaparece, estes grânulos são eliminados. Esta eliminação é feita através de autofagia, mais especificamente por macroautofagia, na qual a proteína *Valosin-containing protein* (VCP) é fundamental, ou por autofagia mediada por chaperonas.

Embora por um lado os grânulos de stress sejam essenciais na resposta da célula ao stresse, por outro lado, defeitos na sua formação, dinâmica ou eliminação podem contribuir para a patogénese de diversas doenças humanas, como por exemplo, cancro, e doenças neurodegenerativas.

De facto, existem cada vez mais evidências que ligam os grânulos de stresse a doenças neurodegenerativas, nomeadamente as doenças de Alzheimer, Parkinson, Huntigton ou Esclerose Lateral Amiotrófica. Para além disso, várias das proteínas envolvidas na patogénese destas doenças também se encontram presentes nos grânulos de stress, como é o caso da Atx-2, Tau, FUS ou TDP-43. A importante função dos grânulos de stresse a

nível celular e a sua implicação em diversas patologias torna-os alvos importantes para o estudos dessas doenças e até como possíveis alvos terapêuticos.

O principal objetivo deste trabalho foi identificar todos os componentes proteícos dos grânulos de stresse descritos em células de mamífero, recolher informações sobre os mesmos e catalogar toda a informação recolhida numa base de dados online.

Os componentes proteícos dos grânulos de stresse foram identificados através duma pesquisa exaustiva nos estudos disponíveis no PubMed, tendo sido filtradas apenas as proteínas existentes em células de mamífero. Para cada proteína, informações como o nome completo da proteína, gene ID, UniProt ID, se é ou não uma RBP, função molecular ou se está envolvida nalguma doença, foram obtidas através de bases de dados públicas. Para além disso, também foram recolhidas informações acerca do tipo celular em que as proteínas foram estudadas e tipo de estímulo utilizado para induzir o stress.

Uma análise transcriptómica foi efetuada, utilizando dados de estudos que analisaram a expressão génica em indíviduos com as doenças de Alzheimer, Parkinson, Huntington, Esclerose Lateral Amiotrófica e também em indíviduos não-doentes (controlos). Nesta análise os grupos de doença *versus* controlo foram comparados, de forma a investigar a expressão diferencial dos componentes dos grânulos de stress.

Foram catalogadas 464 proteínas, das quais 253 foram identificadas como RBPs e 111 já tinham sido previamente ligadas a doenças. Para além disso, 32 proteínas foram identificadas como estando relacionadas com a autofagia. Do total das 464 proteínas identificadas, 225 têm como função molecular a ligação, 308 encontram-se envolvidas em processos celulares e 131 pertencem à classe proteíca de ligação a ácidos nucleícos.

Os dados da análise da expressão génica entre os grupos "doença" e "controlo", demonstraram que a grande maioria dos componentes dos grânulos de stress encontram-se com a expressão alterada nas doenças neurodegenerativas estudadas. Por exemplo, verificou-se que na doença de Alzheimer, 187 componentes tinham a sua expressão significativamente aumentada, enquanto que 193 tinham a sua expressão diminuída.

O perfil proteíco dos grânulos de stresse varia consoante o tipo celular ou tipo de estímulo de stress induzido. Assim, foi possível observar que a maioria das proteínas, mais especificamente 327, foram identificadas nas células U-2OS, enquanto 427 foram identificadas em estudos que utilizaram o arsenito de sódio como estímulo.

Toda a informação recolhida foi catalogada na base de dados *Mammalian Stress Granules Proteome* (MSGP), disponível para todos os investigadores em https://msgp.pt/. A base de dados MSGP possui a listagem de todos os componentes proteícos dos grânulos de streses assim como informações dos mesmos e dados de expressão génica para cada uma dos componentes identificados. O utilizador pode pesquisar por uma proteína específica utilizando a "*search bar*" ou pode filtrar a pesquisa utilizando as diferentes categorias e *tags* disponíveis.

A MSGP é a primeira base de dados que reune todos os componentes dos grânulos descritos, constituindo uma importante ferramenta para os investigadores desta área, tendo para além disso, a possibilidade de continuar em crescimento, através da adição de informações não só acerca dos componentes já catalogados, como de outros componentes dos grânulos de stress, como por exemplo microRNAs.

Palavras-chave: Stresse; agregação; mamífero; componentes, proteínas, neurodegenerescência; base de dados; expressão génica

INDEX OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

INDEX OF ANNEXES

LIST OF ABBREVIATIONS

AD - Alzheimer's disease

ALS - Amyotrophic Lateral Sclerosis

Atx-2 - Ataxin-2

CMA - Chaperone-mediated autophagy

CNV - Copy Number Variations

CPEB - Cytoplasmic polyadenylation element-binding protein 1

DLB - Dementia with Lewy Bodies

DLPFC - Dorsolateral prefrontal cortex

dsRBM - double-stranded RNA-binding motif

eIF2α - Eukaryotic Initiation Factor 2 α

EIF2AK1 - Translation Initiation Factor 2 Alpha Kinase 1

EIF2AK3 - Eukaryotic Translation Initiation Factor 2 Alpha Kinase 3

EIF2AK4 - eIF2α kinase 4

eIF4A - Eukaryotic translation initiation factor 4A

eIF4E - Eukaryotic translation initiation factor 4E

eIF4F - Eukaryotic initiation factor 4F

eIF4G - Eukaryotic translation initiation factor 4G

ERs - Estrogen receptors

FC - fold-change

FMRP - Fragile X mental retardation protein

FTD - Frontotemporal Dementia

FXS - Fragile-X syndrome

G3BP1 - Ras GTPase-activating protein-binding protein 1

GCN2 - General control non-derepressible-2

H_2O_2 - Hydrogen peroxide

HD – Huntington's disease

HIV - Human Immunodeficiency Virus

hnRNPA1 - Heterogeneous Nuclear Ribonucleoprotein A1

hnRNPA2 - Heterogeneous Nuclear Ribonucleoprotein A2

HRI - Haem-regulated inhibitorEukaryotic

hsp70 - 70 kilodalton heat shock proteins

IDDs - Intrinsically disordered domains

IRES - Internal ribosome entry site

ISR - Integrated stress response

ITAFs - IRES trans-acting factors

Met-tRNA$_i^{Met}$ - Met-charged initiator tRNA

mRNA – Messenger RNA

mRNPs - Messenger ribonucleoproteins

MSGP - Mammalian Stress Granules Proteome

NCBI - National Center for Biotechnology Information

NF-κB - Nuclear factor κB

OH - Hydroxyl radical

OMIM - Online Mendelian Inheritance in Man

PABP1 - Polyadenylate-binding protein 1

PANTHER – Protein Analysis Through Evolutionary Relationships

P-Bodies – Processing Bodies

PD – Parkinson's disease

PEK - Pancreatic eIF2α kinase

PERK - PKR-like ER kinase

PKR - Protein kinase R

PLDs - Prion-like domains

RBPs - RNA-binding proteins

RNAP2 - RNA polymerase II

RNAP3 - RNA polymerase III

RNPs - Ribonucleoproteins

ROS - Reactive Oxygen Species

RRM - RNA-recognition motif

SGD - *Saccharomyces* Genome Database

SGs - Stress Granules

SMN - Survival motor neuron protein

SRC3 - Steroid receptor coactivator 3

TIA-1 - T-cell intracellular antigen-1

TIAR - TIA-1-related

TNF - Tumor Necrosis Factor

TRAF2 - TNF Receptor-associated Factor 2

TTP - Tristetraprolin

VCP - Valosin-containing protein

ZBP1 - Z-DNA-binding protein 1

CHAPTER 1- INTRODUCTION

1. Introduction

1.1. The cellular response to stress

Eukaryotes cells are subject to stressful conditions that can disrupt their normal functioning and lead to a conservative stress response, which is needed to overcome that stress and restore cellular homeostasis. Conditions such as heat, nutrient deprivation, oxidative state, or genotoxic and osmotic stress the stress stimuli known to lead to the assembly of stress granules (SGs), which are essential players in the cellular response to stress (Ghisolfi, Dutt, McConkey, Ebert, & Anderson, 2012)(Pothof, Verkaik, Hoeijmakers, & Van Gent, 2009)(Souquere et al., 2009). These irregularly shaped structures with no membranous border, have a moderate electron density and its size can go from one to several micrometers. Stress granules are transiently formed with the main objective of reestablishing cellular homeostasis during a stress response (Ghisolfi et al., 2012)(Souquere et al., 2009).

Along with SGs, stress proteins are responsible for the primary mediation of cellular responses to stress. Besides being involved in stress response, these proteins are also involved in normal life processes, thus being essential for cell survival (Milisav, 2011). Depending on the stressor, the stress response can be regulated in different ways. Nonetheless, in general, the stress response is initiated due to the detection of denatured proteins (Milisav, 2011)(Poljšak & Milisav, 2012). According to the Mosby's Medical Dictionary (10th edition), denatured proteins are proteins that due to some stress stimuli, such as heat and radiation, or certain compounds, such as organic solvents like alcohol, lose their secondary, tertiary and quaternary structure. The loss of the proteins' 3D structure is responsible for their loss of function, which consequently may lead to the disruption of cell activity and cell death (Samson et al., 2016). In this loss of structure, is particularly important the hydrogen bonds are fundamental for proteins to acquire the right structure, which is fundamental for their function. These bonds, are sensitive to stressors like heat, which leads to protein denaturation and the respective cellular response to stress (Samson et al., 2012).

1.1.1 Heat

In the case of heat-induced stress, high temperatures lead to protein denaturation and there is a consequent stress response without the need for specialized thermosensitive proteins

(Milisav, 2011)(Poljšak & Milisav, 2012). For this specific response, however, a certain class of stress proteins is necessary: the heat shock proteins. Since with high temperatures, there is an increase of denatured proteins, heat shock proteins act as chaperones in protein folding, making sure proteins maintain their structure and therefore, their function (Richter, Haslbeck, & Buchner, 2010).

1.1.2. Oxidative stress

Oxidative stress affects the normal function of cells and damages important cellular components such as DNA, proteins, and lipids. This type of stress occurs due to an imbalance between the production of reactive oxygen species (ROS), such as H_2O_2 (hydrogen peroxide) and OH (hydroxyl radical) and their detoxification (Valko et al., 2007)(Joseph, Zhang-James, Perl, & Faraone, 2015). Although ROS are important for the immune response against pathogens, oxidative stress has been shown to be involved in a variety of human diseases (Segal, 2005). For example, several studies point to an involvement of oxidative stress in different neurodegenerative diseases (Patel & Chu, 2011), including Alzheimer's disease (Valko et al., 2007), Parkinson's disease (Article & Hwang, 2013), Huntington's disease and Multiple sclerosis (Haider et al., 2011). Oxidative stress has been also linked to cancer (Halliwell, 2007), autism (James et al., 2004) and heart problems such as myocardial infarction (Ramond et al., 2011)(Dean et al., 2011) and heart failure (Singh, Dhalla, Seneviratne, & Singal, 1995), amongst other diseases.

1.1.3. Genotoxic stress

Genotoxic stress is responsible for the damage of genetic information in a cell. This damage can cause mutations and therefore lead to different diseases such as cancer (Coates, Lorimore, & Wright, 2005).

1.1.4. Osmotic stress

Osmotic stress is caused by a change in solute concentration around the cell. This change affects the movement of water across the cell membrane and interferes with the transport of substrates (Jolla, 2006). To counteract this stress, the cell has signals that provide information about the osmolarity of its surroundings, which allows the cell to activate a response in the case of extreme conditions (Kültz & Burg, 1998)(Ku, 2007). Like for the

previous types of stress, it has been shown that osmotic stress seems to be related to various disorders, such as inflammatory disorders, for example (Brocker, Thompson, & Vasiliou, 2012).

1.2. The stress granules composition

SGs components include several proteins and mRNAs stalled in translation. In more detailed look for SGs composition, one can find polyA-RNA, poly ADP-ribose (Leung, Todorova, Ando, & Chang, 2012), long non-coding RNAs (Pothof et al., 2009), ubiquitin (M. G. Thomas, Loschi, Desbats, & Boccaccio, 2011), ubiquitin modifying enzymes glucosyltransferases, microRNAs, nuclear transport factors (Mahboubi, Seganathy, Kong, & Stochaj, 2013), components of the small ribosomal subunit, translation initiation factors (M. G. Thomas et al., 2011), signaling molecules (Kedersha, Ivanov, & Anderson, 2013), proteins involved in the regulation of messenger RNA (mRNA) processing, transport and stability, and several enzymes like helicases, phosphatases, GTPases and ribonucleases (Lewitzky, Simister, & Feller, 2012). Despite all this diversity, several proteins are always present in the SGs, constituting their makers, such as Polyadenylate-binding protein 1 (PABP1), Ras GTPase-activating protein-binding protein 1 (G3BP1) or T-cell intracellular antigen-1 (TIA-1). However, exact the protein profile of SGs varies, depending on the cell type and on the type of stress used to induce SGs (M. G. Thomas et al., 2011)(Aulas & Vande Velde, 2015)(J. R. Buchan, Yoon, & Parker, 2011).

1.3. The stress granules formation

1.3.1. The translation process

In eukaryotes, the translation process leads to the formation of proteins from the mRNA. This process occurs in three main phases (Figure 1.1): initiation, elongation, and termination (Sonenberg & Hinnebusch, 2009). The initiation factors interact with the 5'-end of an mRNA molecule, the 5'cap, and with the 5' UTR (Pisareva, Pisarev, Komar, Hellen, & Pestova, 2008)(Sonenberg & Hinnebusch, 2009) (Aitken & Lorsch, 2012). These factors are responsible for the binding to the 40S ribosomal subunit, also known as the small ribosomal subunit (Aitken & Lorsch, 2012). In this phase, it is very important

that the 60S ribosomal subunit (or large ribosomal subunit) does not bind prematurely to the mRNA. For that reason, the eIF3, which also interacts with the eIF4F complex, is associated with the 40S subunit. All of these components form the 43S preinitiation complex or 43S PIC (Aitken & Lorsch, 2012). This complex, together with protein factors, scans the mRNA towards its 3'-end until it reaches the AUG codon, which is the start codon in eukaryotes, encoding methionine (Sonenberg & Hinnebusch, 2009). Then, the eIF2 brings the Met-charged initiator tRNA (Met-tRNA$_i^{Met}$) to the P-site of the small ribosomal subunit. The hydrolyzation of GTP by eIF2 leads to the dissociation of factors from the small subunit and then the large subunit is able to associate with the small one (Pisareva et al., 2008).

The normal translation process, is, however, perturbed during cellular stress, namely by its reduction. As several proteins are needed to the response to stress, those specific mRNAs that need to be translated, do it through what is called "Cap-independent initiation". This translation occurs due to the presence of an Internal ribosome entry site, known as IRES. As the name indicates, in these cases, the 5'cap is not necessary for the initiation phase and there is no need for the scanning of the 5' UTR since the ribosome goes to the start site via direct binding. Nevertheless, initiation factors and/or IRES trans-acting factors (ITAFs) are still needed for this type of translation occurs (López-Lastra, Rivas, & Barría, 2005)(Malys & McCarthy, 2011).

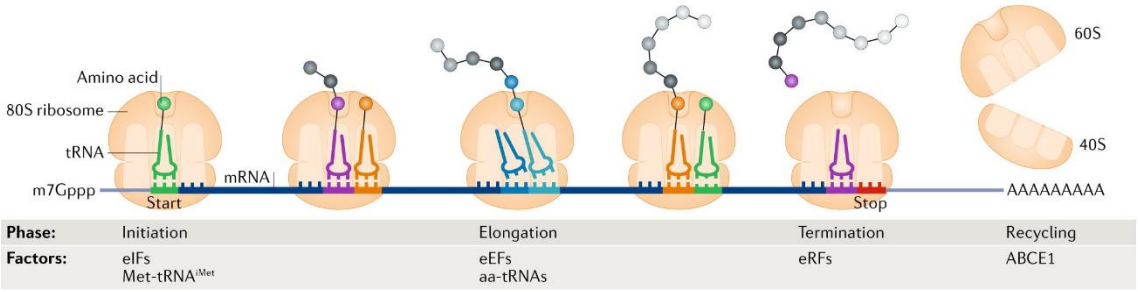


Figure 1.1 Translation phases. The initiation of translation begins when the translation machinery come together at the start codon for the protein synthesis to start. During elongation, the ribosome translocates to the next codon so that the peptide chain can be synthesized. When the ribosome reaches the stop codon the complete peptidic chain is released and the ribosomes are recycled for the nex translation round. eIFs- eukaryotic translation initiation factors; Met-tRNA^{iMet} - initiator methionyl-tRNA; eEFs- eukaryotic elongation factors; aa-tRNAs- aminoacyl-tRNAs; eRFs- eukaryotic peptide chain release factors; ABCE1- ATP-binding cassette sub-family E member 1. (Source: Schuller, A. P.,*et al*, 2018)

The ribosomes are composed of three tRNA-binding sites: the Aminoacyl-tRNA binding, or A-site, the Peptidyl-tRNA binding, or P-site and the Exit site, or E-site (Hellen & Sarnow, 2001)(Hinnebusch, 2017). During the elongation phase of translation, for the next codon to be translated, the ribosome has to translocate to the next mRNA codon, therefore creating an amino acid chain. For the amino acid to be added, when the initiator tRNA occupies the P-site, the aminoacyl-tRNA is positioned in the A-site. Then a peptide bond is formed by peptidyl transferase and the ribosome translocates to the next codon again (Hinnebusch, 2017). Finally, the translation process ends when the ribosome is disassembled and the eukaryotic release factor eRF1, which is able to recognize the eukaryotic stop codons (UAG, UAA, UGA), with the help of eRF3 (a ribosome-dependent GTPase), releases the polypeptide (Schueren & Thoms, 2016).

1.3.2. The eIF2α phosphorylation

When a cell is subjected to a stress stimulus, the serine 51 of Eukaryotic Initiation Factor 2 α (eIF2α) is phosphorylated. The recycling of this kinase to its active GTP-bound form is disrupted and with the reduction of the eIF2– GTP levels, there is a consequent reduction of the translation process. This reduction occurs due to the fact that eIF2– GTP is a part of the eIF2/ $tRNA_i^{Met}$ /GTP ternary complex, which is necessary for the formation of the 48S preinitiation complex, which assembles at the 5′-ends of capped mRNAs (Kedersha et al., 2002). Since the phosphorylation of eIF2α depletes the ternary complex, the formation of the 48S preinitiation complex is disrupted. Therefore, a translationally-stalled, non-canonical 48S complex is produced and it is unable to recruit the 60S ribosomal subunit, affecting translation (Kedersha et al., 2002)(Kimball, Horetsky, Ron, Jefferson, & Harding, 2003). However, the eukaryotic cells are able to reprogram their translational machinery, allowing a selective expression of the proteins necessary for viability. So, the mRNAs that encode "housekeeping" proteins are redirected from polysomes to SGs (Anderson & Kedersha, 2002b). In the absence of eIF2–GTP–$tRNA^{Met}$ and although mRNA is directed to sites of reinitiation, it ends up back in stress granules, where it accumulates (Kedersha et al., 2000). These mRNAs stored in SGs are not degraded, which then can be rapidly reinitiated and used for the translation when cells recover from stress.

1.3.2.1. The four kinases

Besides the reduction of translation, the stress signals also activate gene expression programmes with the purpose of controlling cellular damage or to induce apoptosis, if the cell fails to overcome the stress. There are four kinases crucial to stress adaptation and response, which phosphorylate the α subunit of eIF2 (Table 1.1): General control non-derepressible-2 (GCN2) (or eIF2α kinase 4 (EIF2AK4)), Pancreatic eIF2α kinase (PEK) (or (PKR-like ER kinase (PERK) or Eukaryotic Translation Initiation Factor 2 Alpha Kinase 3 (EIF2AK3)), Protein kinase R (PKR) and Haem-regulated inhibitor (HRI) (or Eukaryotic Translation Initiation Factor 2 Alpha Kinase 1 (EIF2AK1)). These kinases possess specific regulatory regions that recognize different stress conditions, being therefore specifically activated according to the stimulus. (R. C. Wek, Jiang, & Anthony, 2006)

Table 1.1 The four kinases. The functions and stress stimulus activators of the kinases responsible for eIF2α phosphorylation.

	Kinase name	Function	Activated by
GCN2	General control non-derepressible-2	Monitoring amino acid levels (S. A. Wek, Zhu, & Wek, 1995)	Amino acid deprivation, UV irradiation and proteasome inhibition (R. C. Wek et al., 2006)
PEK	Pancreatic eIF2α kinase	Endoplasmic reticulum protein sensor	ER stress, more specifically to misfolded proteins (Harding et al., 2000)(Kaufman, 2004)
PKR	Protein kinase R	Double-stranded RNA-dependent kinase, associated with the anti-viral defense mechanism mediated by interferon (Barber, 2005)	Viral infection, heat and UV irradiation (S. A. Wek et al., 1995)
HRI	Haem-regulated inhibitor	Ensures the balanced synthesis of globin chains and heme during erythrocyte maturation (Ghisolfi et al., 2012)(Lu, Han, & Chen, 2001)	Heat, oxidative stresses (such as sodium arsenite (McEwen et al., 2005)) and haem deprivation in erythroid tissues (Lu et al., 2001)

1.3.2.2. The eIF2α phosphorylation activates NF-κB during stress

Another important factor in the stress response is the Nuclear factor κB (NF-κB), which is activated by the phosphorylation of eIF2α. In response to different types of stress, such as amino acid depletion and UV irradiation, NF-κB coordinates the transcription of a variety of genes involved in inflammatory and immune responses, cell growth and apoptosis (Karin & Ben-neriah, 2000)(Q. Li & Verma, 2002)(Pahl, 1999). In stressful conditions, the eIF2 kinases facilitate the NF-κB activity. For example, when responding to UV irradiation or amino acid starvation, GCN2 phosphorylation of eIF2 is necessary for the activation of NF-κB (H.-Y. Jiang & Wek, 2005)(H. Jiang et al., 2003). On the other hand, PEK enhances NF-κB transcriptional activity in response to ER stress (H. Jiang et al., 2003).

1.3.3. eIF2α- independent stress granules assembly

Although SGs assembly is majority eIF2α-dependent, they can also be formed independently of this factor, in a non-canonic pathway. In this pathway, the steps downstream from eIF2α are inhibited, and it depends on the Eukaryotic initiation factor 4F (eIF4F) complex.

As mentioned, the eIF2/ $tRNA_i^{Met}$ /GTP ternary complex associates with the 40S ribosome, which consequently leads to the formation of the 43S preinitiation complex. The 43S complex (which the eIF2/ $tRNA_i^{Met}$ /GTP ternary complex is then a part of (together with 40S particle and eIF3)), is recruited to mRNA by the eIF4F complex, during cap-dependent translation (Dang et al., 2006). This leads to the formation of the 48S initiation complex.

So, the main difference between the canonical and non-canonical pathway is the mode of induction of SG assembly. The eIF4F cap-binding complex, which is composed by eIF4A, eIF4E, and eIF4G represents a control point for translation initiation. (Pelletier, Graff, Ruggero, & Sonenberg, 2015) Therefore, translation initiation can be affected without eIF2α phosphorylation being required, and changes in the composition and activity of the eIF4F complex can lead to the generation of SGs (Mokas et al., 2009)(Farny, Kedersha, & Silver, 2009)(Kedersha et al., 2013). By compromising the

activity of the eIF4F complex, the translation initiation events can be prevented and consequently, SGs are generated. For example, pateamine A or hippuristanol can be used to alter the helicase activity of eIF4A. (Dang et al., 2006)(Rachid Mazroui et al., 2006)(Cencic et al., 2009) Also, eIF4E can be modulated with hydrogen peroxide or selenite and eIF4G can be cleaved, all of which can lead to a non-canonical SG assembly (eIF2α- independent SGs) (Rachid Mazroui et al., 2006)(Tisdale et al., 2010)(Fujimura, Sasaki, & Anderson, 2012).

1.3.3.1. The eIF4F complex

The eIF4F complex is constituted by the cap-binding protein Eukaryotic translation initiation factor 4E (eIF4E), the scaffolding protein Eukaryotic translation initiation factor 4G (eIF4G), and by the DEAD-box RNA helicase Eukaryotic translation initiation factor 4A (eIF4A) (Mokas et al., 2009)(Cencic et al., 2009). Contrary to eIF4E, the eIF4A is the most abundant initiation factor, although just a small of its proportion is present as an eIF4F subunit (Galicia-Vázquez, Cencic, Robert, Agenor, & Pelletier, 2012)(Grifo, Tahara, Morgan, Shatkin, & Merrick, 1983)(Edery et al., 1983)(A. Thomas, Goumans, Amesz, Benne, & Voorma, 1979). Also, this factor unwinds the mRNA structure, whilst eIF4E is able to bind to the cap structure present at the end of mRNAs (Pestova et al., 2001). The eIF4G interacts with eIF4E and eIF4A through defined domains, provides the scaffold upon which other factors important for the initiation process assemble (Kapp & Lorsch, 2004) and is involved in the recruitment of the 43S pre-initiation complex by interacting with 40S-associated eIF3 (Pestova et al., 2001). There are certain drugs and lipid mediators, such as hippuristanol and pateamine A (Dang et al., 2006)(Rachid Mazroui et al., 2006), which are able to target eIF4A, more specifically its RNA helicase activity (Dang et al., 2006) and consequently are able to inhibit translation initiation.

1.3.3.1.2. The eIF4G leads to NF-kB inhibition

In stressed cells, the eIF4G recruits the adaptor protein TNF Receptor-associated Factor 2 (TRAF2) to SGs. This protein is responsible for linking the Tumor Necrosis Factor (TNF) receptor to a signaling cascade that activates NF-kB (Kim, Back, Kim, Ryu, &

Jang, 2005). By being recruited to SGs, this cascade is inhibited, consequently leading to the repression of the activation of the NF-kB transcription factor (Kedersha et al., 2013), therefore affecting stress response.

1.4. The stress granules assembly process

1.4.1. Stress stimulus

In vitro, several stressors may be used to induce SG assembly. Some of those include sodium arsenite (oxidative stress), heat shock, UV radiation, MG132 (proteasome inhibition), hippuristanol (inhibition of eIF4A), pateamine A (inhibition of eIF4A), thapsigargin (endoplasmic reticulum stress) and puromycin (disassembled polysomes) (Aulas et al., 2017).

Sodium arsenite was the first stressor found to induce SGs and it is considered the most effective agent used to induce SG assembly (Kedersha, Gupta, Li, Miller, & Anderson, 1999). This chemical leads to SG assembly through the canonical pathway, since it activates the HRI kinase which leads to eIF2α phosphorylation and consequent SG formation (Kedersha et al., 2000).

Arsenicals have the ability to bind to reduced cysteines in regulatory proteins and therefore inactivate a great variety of enzymes, such as key anti-oxidant enzymes. Given that arsenite is able to inactivate the protective antioxidant systems that counteract the imbalance caused by ROS, there's an accumulation of ROS. (Shen, Li, Cullen, Weinfeld, & Chris Le, 2013)

Another commonly used stressor is heat shock. The increase of temperature leads to a disruption in protein homeostasis, affecting protein folding and consequently very important cellular processes (Richter et al., 2010)(Balchin, Hayer-Hartl, & Hartl, 2016). Heat stress will lead to protein misfolding, making proteins to lose their configuration and negatively affecting their function (Richter et al., 2010). Also, misfolded proteins can lead to protein aggregation and eventually to plaque formation, an abnormality that can be seen in various neurodegenerative diseases, such as Alzheimer's, Parkinson's and Huntington's disease (Wyttenbach & Arrigo, 2009).

1.4.2. Granule nucleating proteins and primary aggregation

As mentioned the canonical pathway for SGs assembly starts with the eIF2α phosphorylation and the consequent translation arrest. This leads to the aggregation of the stress granule nucleating proteins in the cytoplasm, which will generate the initial foci (Bounedjah et al., 2014) (Figure 1.2). These nucleating proteins are RNA-binding proteins (RBPs) with intrinsically disordered domains (IDDs) or prion-like domains (PLDs). IDDs and PLDs consist of low complexity sequences that support protein aggregation (Gilks et al., 2004), through electrostatic interaction (Lin, Protter, Rosen, & Parker, 2015). There are certain RBPs essential for SGs assembly, due to their ability to associate with untranslated mRNA and their IDDs and PLDs, which can act as scaffolds to facilitate the recruitment of other proteins (Gilks et al., 2004), and also constitute SGs markers. Some of these proteins include the Ras GTPase-activating protein-binding protein 1 (G3BP1), the tristetraprolin (TTP), the T-cell intracellular antigen-1 (TIA-1), the TIA-1-related (TIAR), the fragile X mental retardation protein (FMRP)**,** the Cytoplasmic polyadenylation element-binding protein 1 (CPEB), the Survival motor neuron protein (SMN) and importantly the Polyadenylate-binding protein 1 (PABP1) (Kedersha et al., 2000)(Tourrière et al., 2003)(R. Mazroui, 2002)(Stoecklin et al., 2004). These proteins initiate the stress granule aggregation, mediating the 'primary aggregation'. They not only induce SG assembly and become part of them, but also require nonpolysomal 48S messenger ribonucleoproteins (mRNPs) to do so (McEwen et al., 2005).

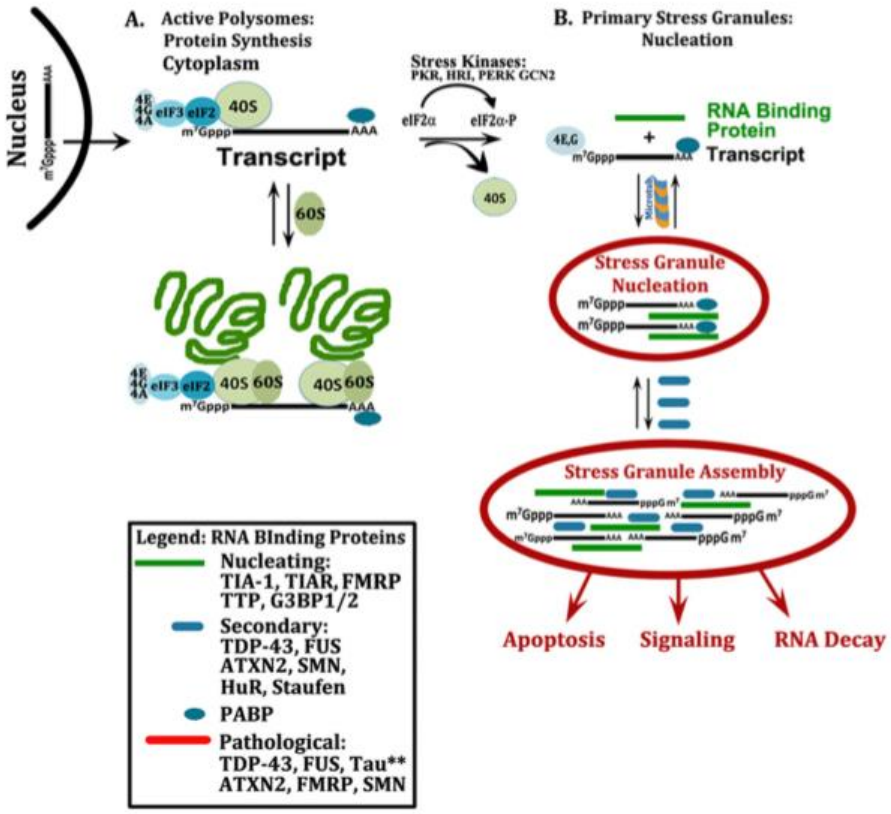


Figure 1.2 Stress granules assembly. A) The eIF4F pre-initiation interacts with the 40S subunit. Protein synthesis begins with the association of this complex with the 60S. **B)** eIF2α is phosphorylated and nucleating RNA binding proteins (RBPs) bind to RNA transcripts. Also, protein-protein interactions occur, initiating stress granule assembly. RBPs then bind to mRNA and other RBPs, increasing the granules complexity. (Source: Wolozin, B., 2012).

1.4.2.1. Secondary aggregation

There are several studies, which showed that the assembly of SGs begins with the formation of small SGs, that then transform into larger granules (Kedersha et al., 2000). In these secondary aggregation events, SGs nucleating proteins are responsible for the connection of the different SGs components and are several types of interactions that promote the growth of the granules (Anderson & Kedersha, 2008)(Kedersha et al., 1999). It is through these protein-protein interactions that non-RNA-binding proteins are also recruited to SGs (Anderson & Kedersha, 2008). Although it is not completely understood, some studies suggest that the rate of SGs assembly can eventually be controlled. For example, several of the SGs components, such as FAST, TIA-1 and TIAR and Steroid receptor coactivator 3 (SRC3) (Kedersha et al., 1999)(W. Li, Simarro,

Kedersha, & Anderson, 2004)(Yu et al., 2007), are nuclear-shuttling proteins, meaning that the regulation of those proteins could eventually control the rate of SGs assembly.

1.4.2.2. The stress granules core and shell

The different steps in the stress granules assembly and the presence of nucleating proteins led to the suggestion that SGs have a stable core and a more dynamic shell of components. In fact, *in vivo* results point to this idea, being the core formed through the oligomerization of mRNPs, whilst the shell formation seems dependent on the IDDs of core components for the recruitment of additional components (Wheeler, Matheny, Jain, Abrisch, & Parker, 2016) (Mahboubi & Stochaj, 2017).

1.5. The integrated stress response

The stress granules formation is part of the cellular integrated stress response (ISR), along with the translational arrest and the polysome disassembly. These events, allow the cell to reprogram its translational repertory through mRNA triage and thus adequate its response to the stress stimulus (Anderson & Kedersha, 2008). Although SGs are transiently expressed, they are able to persist from minutes to hours and once formed they have the ability to intercept several signaling molecules and affect other signaling pathways, which modulate growth, metabolism, and survival.

1.5.1. The mRNA triage model

SGs used to be viewed only as ‘repositories’ for untranslated mRNAs. Those mRNAs, which accumulated during stressful conditions, were, therefore ‘protected’, and when conditions improved they were again translated. However, several studies ended up showing that SGs were more than that (Anderson & Kedersha, 2008). The mRNA triage model establishes SGs as ‘triage compartments’, where specific mRNA transcripts are remodeled, selected for decay, exported or selected for storage, in or out of SGs (Anderson & Kedersha, 2002a)(Anderson & Kedersha, 2006). It is now known that specific RBPs, such as Z-DNA-binding protein 1 (ZBP1), are able to regulate the movement of specific mRNAs in SGs, thus contributing to the control of the mRNA fate

through that regulation (Stöhr et al., 2006). During stress this protein is recruited to SGs, bound to its associated mRNAs and is able to retain and enhance the stability of those transcripts within SGs.

1.6. The stress granules functions

The SGs assembly is a process of regulated protein aggregation that has several advantages for cell physiology, as it helps in the control of proteostasis and ribostasis: a) its formation is a more rapid than transcriptional or translation alterations, b) its disassembly upon stress relief allows the cell to have different proteins and factors ready to use, and c) important cellular factors are protected from degradation upon SG assembly (Arimoto, Fukuda, Imajoh-Ohmi, Saito, & Takekawa, 2008)(Panas, Ivanov, & Anderson, 2016). For example, by sequestering signaling molecules and regulating mRNA translation, SGs are able to protect the cell from apoptosis (Mahboubi & Stochaj, 2017).

Under stress conditions and upon phosphorylation of eIF2α, the translation initiation is arrested, global protein translation in reduced and SG are assembled. In line with these events, it was hypothesized that SG were foci where mRNAs translation was repressed (Anderson & Kedersha, 2008). Indeed, several SG components are translational repressors, and SG formation is positively correlated with a decrease in the global translation levels (Kedersha & Anderson, 2009). Moreover, the assembly of proteins and mRNAs to SG might limit the availability of those components in the translation machinery, thus contributing to a repression effect (Anderson & Kedersha, 2008)(Kedersha & Anderson, 2009). In fact, several studies reported that alterations in SG components (mainly RBPs) lead to a repression (not complete, however) of specific mRNAs (Kedersha et al., 2000)(Moeller, Cao, Li, & Dewhirst, 2004). For example, the ablation of ataxin-2 led to a reduction in the global translation rate (Fittschen et al., 2015), while the knockdown of XRN1 is impaired with the translational repression, triggered by NMDA in neurons (Luchelli, Thomas, & Boccaccio, 2015). Besides RBPs, in SG are also assembled other translation players such as small ribosomal subunits, translation initiation factors and different signaling molecules (Kedersha & Anderson, 2009), strengthening the idea that SG might be important and active players in translational repression. Moreover, it was suggested that SG might be integrated with miRNA-induced translational silencing pathways (Emde & Hornstein, 2014). Despite all these evidences,

other studies suggest that SGs formation is not essential for a global translation repression (J. Ross Buchan, Muhlrad, & Parker, 2008)(Mokas et al., 2009)(Loschi, Leishman, Berardone, & Boccaccio, 2009) or that the impairment of SGs assembly by depletion of core factors did not affect global protein synthesis (Ohn, Kedersha, Hickman, Tisdale, & Anderson, 2008)(Mokas et al., 2009). Thus, it is not completely clear if SGs assembly per se is important for the translation repression of certain mRNAs and further studies are needed to establish SGs as definitive players in these mechanisms.

A contrasting function for SG proposes that their formation might promote the assembly of translation complexes (J Ross Buchan & Parker, 2009), due to the local increase in mRNAs and translation factors, which might facilitate the translation initiation. It was reported that upon stress there is a reduction in ~25% of certain mRNAs, whereas other 25% of mRNAs (including heat-shock protein transcripts) are increased (Anderson & Kedersha, 2008). Thus, SG formation might be essential for the translation of mRNAs directly implicated in the stress response, although the absence of 60S subunits in SGs (Anderson & Kedersha, 2002a), suggest that translation itself is unlikely in these foci. It was proposed that SG could act as centers for mRNAs storage and triage, thus influencing protein expression (Kedersha et al., 2000). It was shown that certain mRNAs are recruited to SG during stress, other transcripts are selectively exported to Processing bodies (P-bodies) for degradation, while several mRNAs encoding proteins involved in stress response, like for example heat shock proteins are excluded from SG (Kedersha et al., 2005). Therefore, the SG assembly could allow cells preventing the accumulation of misfolded proteins by reducing the synthesis of certain transcripts while optimizing the translation of mRNAs involved in stress response.

It was also proposed that SG might be implicated in the stabilization of mRNAs, as they contain several stabilizing proteins like HuR or ZBP1 (Rachid Mazroui, Marco, Kaufman, & Gallouzi, 2007)(J Ross Buchan & Parker, 2009). The recruitment of these proteins to SG conditions their cytoplasm availability, thus limiting their function. For example, it was shown that ZBP1 knockdown induced a selective destabilization of target mRNAs (Stöhr et al., 2006). However, some studies oppose to this idea suggesting that SGs are not required for mRNAs stabilization during stress (Hilgers, Teixeira, & Parker, 2006)(J. Ross Buchan et al., 2008)(Bley et al., 2015). A recent study showed that the stabilization of bulk mRNA including IGF2BP1 target transcripts is largely independent of SGs formation (Bley et al., 2015).

SGs assembly and dynamics might be important in the cellular decision of entering or not in apoptosis depending on the response to stress. In fact, to the SGs are recruited several apoptosis regulatory factors, which could inhibit or delay stress-induced cell death signaling (Kedersha et al., 2013). In severe apoptosis-inducing stress, the RACK1 protein binds to the stress-responsive MTK1 kinase and facilitates its activation. However, during modest stress, RACK1 is recruited to SG, thus limiting MTK1 kinase activation and avoiding apoptosis (Arimoto et al., 2008). The recruitment of regulatory-associated protein of mTOR (Raptor) to SG also prevents the over-activation of mTORC1 signaling, thus inhibiting apoptosis (Thedieck et al., 2013). In stressed cells the SG formation reduces the production of reactive oxygen species, thereby preventing apoptosis (M. Takahashi et al., 2013). Moreover, it was shown that upon a cold shock-stress global protein synthesis is suppressed, inducing SG and ensuring cell survival during the stress (Hofmann, Cherkasova, Bankhead, Bukau, & Stoecklin, 2012). Also in line with this function, other studies reported that impairing SG assembly leads to a decrease in cell viability after stress exposure (Baguet et al., 2007)(Kwon, Zhang, & Matthias, 2007). It was also shown that the inhibition of SG formation by oxidizing TIA1, made cells more vulnerable to apoptosis (Arimoto-Matsuzaki, Saito, & Takekawa, 2016). Altogether, these data suggest that SG formation might function as a defense mechanism protecting cells from apoptosis under stress conditions, by regulating mRNA translation and sequestering signaling molecules. This role of SG could also be important in the cellular response to viral infection, which leads to a global protein synthesis suppression and SG assembly (Ruggieri et al., 2012). This importance is highlighted by the fact that several RNA and DNA viruses inhibit the induction of the SG response very soon after infection (Lloyd, 2012), through the sequestration of the SG component G3BP1 by viral proteins (Panas et al., 2015). Altogether these evidence show that SGs are more than storage foci for cells during a stress event, playing instead an active an important role in the stress response.

1.6.1. The stress granules components and its molecular functions

In general, SGs are composed of stalled pre-initiation complexes: 40S subunits, translation initiation factors, poly(A)$^+$ mRNAs and RBPs. Nevertheless, SGs composition is different according to the type of stress and changes during the stress response (T Vanderweyde, Youmans, Liu-Yesucevitz, & Wolozin, 2013). Despite the existence of

several studies reviewing SGs components, to date, there is none detailing all the components identified until now. The proteins that constitute SGs (Table 1.2) are involved in several biological processes and have many functions at a molecular level, which could also provide important hints of further SGs functions.

- ATP binding

According to UniProt (www.uniprot.org), ATP binding proteins are proteins able to bind ATP: ribonucleotide adenosine, which is the source of energy and phosphate for the cell.

- DNA binding

DNA binding proteins have DNA-binding domains. These proteins have an affinity for single or double-stranded DNA. (Pabo & Sauer, 1984)

- Endonuclease activity

Endonucleases are enzymes able to "break" the polynucleotide chain. These proteins catalyze the hydrolysis of ester linkages, or in other words, cleave phosphodiester bonds. (Slor, 1975)

- RNA and mRNA binding

According to the *Saccharomyces* Genome Database (SGD) (www.yeastgenome.org), mRNA binding proteins are able to interact non-covalently with mRNA. RBPs, are proteins able to bind to single or double-stranded RNA. They can be nuclear or cytoplasmic proteins and have very important roles in a variety of cellular functions, such as splicing, mRNA localization, and translation (Lunde, Moore, & Varani, 2007)(Hogan, Riordan, Gerber, Herschlag, & Brown, 2008)(Glisovic, Bachorik, Yong, & Dreyfuss, 2008).

- AU-rich element binding

In mammalian cells, AU-rich elements, also known as Adenylate-uridylate-rich elements, are often determinants of RNA stability (Chen & Shyu, 1995) and can be found in the 3'UTR of many mRNAs. These elements consist of a region with frequent adenine and uridine bases (Shaw & Kamen, 1986)(Barreau, Paillard, & Osborne, 2005). Defects in the function of AU-rich elements may lead to problems in mRNA stability, which have been identified in diseases such as cancer. (Barreau et al., 2005)

- Poly(A) binding

Poly (A)-binding proteins, also known as PABP, are RNA-binding proteins capable of binding to the poly(A) tail on the 3' end of mRNA (Deo, Bonanno, Sonenberg, & Burley, 1999)(Kahvejian, Svitkin, Sukarieh, Boutchou, & Sonenberg, 2005). These proteins have an important role in mRNA metabolism, like in nonsense-mediated decay and take part in the regulation of mRNA production by protecting the poly(A) tail from degradation (Deo et al., 1999)(Gorgoni, 2004).

- Estrogen receptor binding

Estrogen receptors (ERs) are proteins that are activated by the estrogen hormone (17β-estradiol) (Katzenellenbogen et al., 2006). These receptors have different functions, amongst them their function as DNA-binding transcription factors, since they are able to bind to DNA and regulate various genes (Levin, 2005). These receptors are also implicated in disease, more specifically breast cancer since there's an overexpression of these receptors in the majority of cases of "ER-positive" cancers (Fowler & Alarid, 2007).

- Ionotropic glutamate receptor binding

Ionotropic glutamate receptor, also known as iGluRs, are ligand-gated ion channels. They are activated by glutamate and have an important role in synaptic plasticity and in excitatory synaptic transmission (Nath et al., 1988). These receptors can have different ligand binding properties, therefore there are different iGluRs subtypes (AMPA receptors, NMDA receptors, kainate receptors, and delta receptors) (Collingridge, Olsen, Peters, & Spedding, 2009).

- 14-3-3 protein binding

14-3-3 proteins are regulatory molecules able to bind to a great variety of signaling molecules, such as kinases and phosphatases. These proteins can be found in elevated amounts in the cerebrospinal fluid of patients with Creutzfeldt-Jakob disease (H. Takahashi et al., 1999).

- C-C chemokine binding

C-C chemokine binding proteins are proteins able to interact with C-C chemokines. C-C chemokines are signaling proteins that are part of the C-C family of chemokines. (Fernandez & Lolis, 2002)(Zlotnik & Yoshie, 2012) These proteins are responsible for the induction of the migration of monocytes and dendritic cells and are able to induce chemotaxis. (Fernandez & Lolis, 2002)(Le, Zhou, Iribarren, & Wang, 2004)

- Protein kinase binding

A protein kinase is an enzyme able to add phosphate groups on other protein, which may result in a functional change of that protein. (Hanks & Hunter, 2018)

- Dynein complex binding

Proteins with dynein complex binding are able to interact with dynein, a motor protein that moves along microtubules in the cell. (Carter, 2013) Dynein is responsible for the transport of cargo and converts energy stored in ATP to mechanical work. (McKenney, Huynh, Tanenbaum, Bhabha, & Vale, 2014)

- G-quadruplex RNA binding

G-quadruplex RNA are helical structures formed by sequences rich in guanine. (Creacy et al., 2017)

- Ion channel binding

Ion channels are membrane proteins with channel pores, through where ions are able to pass. Since these channels gate the flow of ions across the membrane, they are able to regulate certain electrical signals, such as the membrane potential and action potential, and also regulate the cell volume (Heine, Heck, Ciuraszkiewicz, & Bikbaev, 2019).

- Microtubule binding

Microtubules are dynamic structures that form part of the cytoskeleton. They consist of tubulin polymers and are important for intracellular transport, also providing structure to the cytoplasm of eukaryotic cells. Microtubules are involved in the movement of organelles and in cell division. (Vale, 2003)(Petroni, Jensen, Ladinsky, McDowall, & Pilhofer, 2011)

- Metal ion binding

According to the Mouse Genome Informatics database, metal ion binding proteins are proteins able to interact selectively with any metal ion, such as Li+, Na+, and Mg2+. (Yamashita, Wesson, Eisenmant, & Eisenberg, 1990)

- Ribosome binding

A ribosome is a complex machine that is essential for protein synthesis, being responsible for the linkage of the amino acids. This machine is composed of two components: the small ribosomal subunit and the large ribosomal subunit. (Cruz, Karbstein, & Jr, 2015)

Table 1.2 Selected stress granules components. Some of the SGs proteins and some examples of biological processes to which they are associated and their molecular functions (www.uniprot.org).

	Protein name	Biological processes	Molecular function
G3BP1	Ras GTPase-activating protein-binding protein 1	Ras protein signal transduction	ATP binding, DNA binding, endonuclease activity, mRNA binding, RNA binding
TIA-1	T-cell intracellular antigen-1	Apoptosis	AU-rich element binding, poly(A) binding, RNA binding
TIAR	TIA-1-related	Germ cell development	AU-rich element binding, DNA binding, RNA binding
FUS	Fused in Sarcoma	Cellular response to calcium ion	DNA binding, estrogen receptor binding, identical protein binding, ionotropic glutamate receptor binding
TTP	Tristetraprolin	Cellular response to epidermal growth factor stimulus	14-3-3 protein binding, AU-rich element binding, C-C chemokine binding, DNA binding, protein kinase binding, RNA binding
FMRP	Fragile X mental retardation protein	Cellular response to UV	Chromatin binding, dynein complex binding, G-quadruplex RNA binding, ion channel binding, microtubule binding, mRNA binding
CPEB	Cytoplasmic polyadenylation element-binding protein	Cellular response to hypoxia	Metal ion binding, ribosome binding, translation factor activity, RNA binding
SMN	Survival motor neuron protein	Nervous system development	Identical protein binding, RNA binding

1.7. The stress granules disassembly process

SGs are only transiently formed, disassembling when the stress insult is removed. This disassembly can occur within minutes and is accompanied by the restoration of translation (Kedersha et al., 1999)(Cherkasov et al., 2013). SGs can be dissolved by disaggregation mediated by chaperones or by macroautophagy (Ross Buchan, 2014). For example, several studies have implicated the 70 kilodalton heat shock proteins (hsp70) family in the dissolution mediated by chaperones. The overexpression of hsp70, for example, is able to promote the removal of SGs from stressed cells (Gilks et al., 2004)(Rachid Mazroui et al., 2007).

1.7.1 Autophagy

Autophagy is a regulated cellular process, which allows the cell to degrade dysfunctional or unnecessary components (Klionsky, 2008) (Mizushima & Komatsu, 2011)(Kobayashi, 2015). It is mediated by autophagy-related genes, or ATG, which were first identified in *Saccharomyces cerevisiae* (Tsuboi, 1992)(Klionsky, Cueva, & Yaver, 1992)(Koch et al., 1994). There are three forms of autophagy: macroautophagy, microautophagy and chaperone-mediated autophagy (CMA).

Macroautophagy (Figure 1.3) consists of the intracellular degradation of unnecessary and damaged components or invading microorganisms (Levine, Mizushima, & Virgin, 2011). For this to occur, it is necessary a double layered spherical structure, called autophagosome. This vesicle isolates the targeted constituent from the rest of the cell, travels to a lysosome and fuses with it forming the autolysosome (Mizushima, Ohsumi, & Yoshimori, 2002)(Xie & Klionsky, 2007) (Mizushima, Yoshimori, & Ohsumi, 2011). Then the content is degraded by acidic lysosomal hydrolase, and/or recycled. The specialized SGs disassembly by macroautophagy is also called granulophagy (J Ross Buchan, Kolaitis, Taylor, & Parker, 2013). In this process, SGs are eliminated through their engulfment by autophagic vesicles. For this disaggregation to occur, it is necessary the presence of Valosin-containing protein (VCP), an ATPase essential for autophagy (J Ross Buchan et al., 2013)(Seguin et al., 2014).

In microautophagy, the cytoplasmic cargo is directly engulfed in the lysosome by invagination (Hafner Česen, Pegan, Špes, & Turk, 2012). This specific pathway is

important for cell survival under specific conditions, such as starvation, although there is no evidence for its involvement n SGs disassembly.

CMA is considered a very specific pathway (Levine et al., 2011), where a soluble cytoplasmic protein is recognized and binds to the hsc70-containing complex, resulting in the CMA- substrate/chaperone complex (Bandyopadhyay, Kaushik, Varticovski, & Cuervo, 2008)(Hafner Česen et al., 2012). This complex is targeted to lysosomes and with the help of the lysosomal hsc70 chaperone, the targeted protein is translocated across the lysosome membrane, without the need for formation of additional vesicles (Mizushima et al., 2002)(J. Lee, Giordano, & Zhang, 2011)(Kaushik & Cuervo, 2012). Then, the degradation of the target protein occurs.

It has been demonstrated that the overexpression of the chaperone hsp70 prevents SG formation and its depletion maintains SGs in the cell for a longer period of time (Rachid Mazroui et al., 2007). Furthermore, it has also been established that chaperones, such as hsp27, hsp70, and VCP, which tend to be present in SGs, are required for SG disassembly (Kedersha et al., 1999)(J Ross Buchan et al., 2013)(Seguin et al., 2014)(Hyman et al., 2017). However, it is still not clear the exact molecular mechanisms that connect these chaperones to SG disassembly or the specific proteins that are targeted by these factors. Hence, more studies are needed. (Hyman et al., 2017)

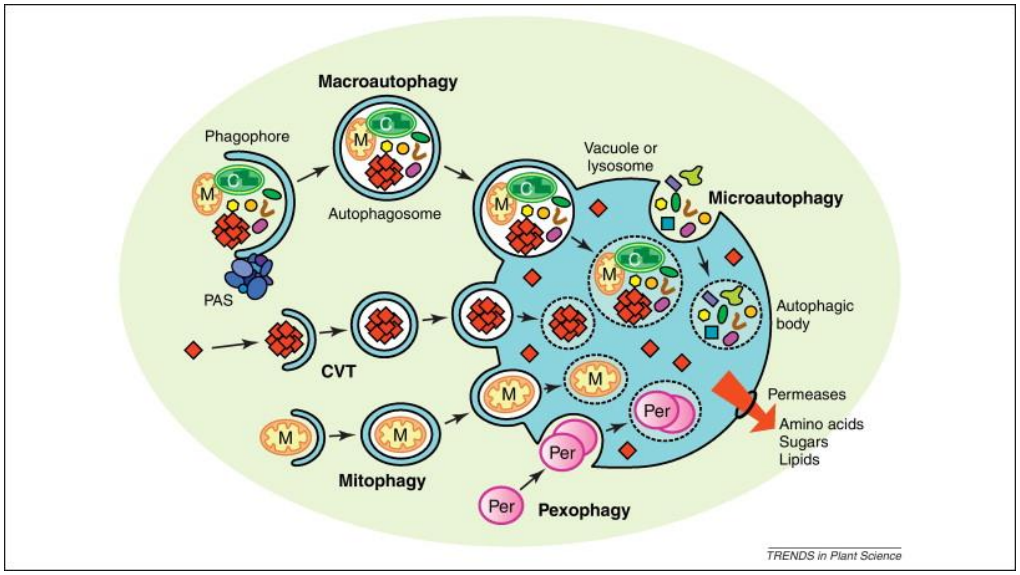


Figure 1.3 The steps during macroautophagy and microautophagy. In macroautophagy, the phagophore is generated, an engulfing membrane that captures cytosolic components and later forms the autophagosome. This double-membrane releases its internal content in the vacuolar lumen and the cargo is then degraded. In microautophagy, the cargo is directly engulfed into autophagic bodies within the vacuole. C- chloroplast; M- mitochondrion; PAS- pre-autophagosomal structure; Per- peroxisome. (Source: Li, F., *et al*, 2012)

1.7.1.1. Autophagy in disease

It is now known that the dysregulation of autophagy is involved in the pathogenesis of a variety of diseases, such as cancer (Moosavi et al., 2018). It was shown that autophagy is not just important for the protection against cancer but also for cancer growth (Iman Tavassoly, 2015). Autophagy is also important in the cellular response to different types of stress, (Paglin et al., 2001), such as nutrient deprivation and hypoxia, as it allows the recycling of ATP and the maintaining of the cellular energy production. On the other hand, it has been shown that when autophagy-related genes are inhibited, there is an increase in cell death, which could be studied as a therapeutic target in oncology (Jin & White, 2007)(Yang, Chee, Huang, & Sinicrope, 2011)(I. Tavassoly et al., 2015).

1.8. Stress granules and disease

As mentioned, RBPs have a wide range of functions being essential in the regulation of gene expression, post-transcriptional processes like pre-mRNA splicing, mRNA cytoplasmic export, turnover, storage, translation, and degradation. Due to these important functions, it is not surprising that a deregulation in the expression of different RBPs seems to underlie a variety of human disorders, including cancer and neurodegenerative diseases (Lukong, Chang, Khandjian, & Richard, 2008)(Cooper, Wan, & Dreyfuss, 2009). For that, it is also easy to conceive that SG might be involved in different human conditions like cancer (Anderson, Kedersha, & Ivanov, 2015), aging (T Vanderweyde et al., 2013), and neurodegenerative disorders (Wolozin, 2012). For example, in human breast cancer biopsies, cancer cells sometimes accumulate MLN51 in discrete cytoplasmic foci resembling SG (Baguet et al., 2007). Interestingly, different studies associated the formation of SG to the resistance to chemotherapeutic drugs (Fournier, Gareau, & Mazroui, 2010)(Adjibade et al., 2015). It was found that at least 30 different RBPs were upregulated in different types of cancers. The authors of this study proposed the idea that fluctuating RBPs levels could result in an increase of non-specific protein interactions with an important impact on the disease outcome (Kechavarzi & Janga, 2014). It was also found that the expression of several RBPs was reduced through aging, suggesting that they could play important roles in maintaining tissue homeostasis with advancing age (Masuda, Marasa, Martindale, Halushka, & Gorospe, 2009). The

implication of RBPs and SG in the pathogenesis of different conditions affecting human health seems credible, however, further studies are needed to establish this link.

1.8.1 Viral infections

When a virus enters a cell, SGs crucial players in the limitation of the viral replication and in the stimulation of the immunologic functions. So, viruses developed several mechanisms to affect SGs formation and also to dissolve the potentially already existing granules. For example, viruses, such as Human Immunodeficiency Virus (HIV), Herpes Simplex, Influenza, are able to affect various stages of SGs biogenesis (Panas et al., 2015). More specifically, certain viruses, like HIV-1, can sequester the protein G3BP1 – an RBP fundamental for SGs nucleation- thus inhibiting the granules' formation. Also, the Dengue virus can capture the TIA-1 protein, another essential protein for SGs assembly. So, the integrity of these granules can be negatively affected by the interaction of viral proteins and these essential SGs components, therefore compromising the cellular response to viral infection (Panas et al., 2015)(Valiente-Echeverría et al., 2016)(Dougherty, Tsai, & Lloyd, 2015)(Poblete-Durán, Prades-Pérez, Vera-Otarola, Soto-Rifo, & Valiente-Echeverría, 2016).

1.8.2. Cancer

During tumor development, the need for nutrient and oxygen increases rapidly and the tumor can easily outgrow the existing vasculature. In this process, cells can be exposed to a great amount of stress, due to nutrient starvation and hypoxia (Ackerman & Simon, 2014). Plus, an imbalance in protein synthesis can cause the endoplasmic reticulum (ER) to overload and consequently lead to ER stress in cancer cells (Clarke, Chambers, Liniker, & Marciniak, 2014). So, cancer cells change their own metabolism to adapt to this environment (Porporato, Dhup, Dadhich, Copetti, & Sonveaux, 2011). But this change leads to an increase in the level of production of ROS and consequently in oxidative stress. ROS are small molecules that result from the malfunctioning of the mitochondria. They are balanced by the action of antioxidants, so when that capacity is exceeded, the stress response is activated, consequently leading to the assembly of SGs (Gorrini, Harris, & Mak, 2013).

SGs can be found in a great variety of tumors, with different histological origins. A few examples are carcinomas, colorectal cancer, and glioblastomas (Vilas-Boas et al., 2016)(Adjibade et al., 2015)(Somasekharan et al., 2015).

It is important to mention that although SGs seems to be prosurvival, they can form as a response to chemotherapy and radiotherapy, ending up compromising the treatment. Different chemotherapeutic drugs often use different signaling pathways to SG formation. For example, selenite-induced SGs assembly is independent of eIF2α phosphorylation and this type of SGs assembly it was shown to promote cell death (Fujimura et al., 2012). On the other hand, cell survival is promoted when 5-FU induce SGs assembly. Due to these opposite effects, it would be very important for cancer therapy to better understand the role of chemotherapeutic drugs in SGs assembly (Anderson et al., 2015)(Moeller et al., 2004).

SGs have also been linked to metastasis. It has been shown that SGs contribute to the enhancement of tumor cell survival during invasion and contribute to the tumoral resistance to chemotherapy, and metastasis (Somasekharan et al., 2015). SGs contribute to the support of cancer cell survival and tumor growth, which turns them into promising targets for cancer therapy.

Several SGs components have also a very important role in cancer, given their capacity to regulate certain cancer-relevant targets. For example, RACK1, a scaffolding protein that interacts with various kinases, phosphatases, apoptosis-related molecules, amongst others, can be found up-regulated in many cancers, such as breast cancers, lung and gastric. Since it has the ability to interact with a great variety of partners, therefore influencing different signaling pathways, it can promote or suppress cancer cell proliferation depending on its interactions. (Adams, Ron, & Kiely, 2011) (J. J. Li & Xie, 2014) Another example is TTP and HuR (ELAV1), which have shared targets, such as cell growth factor, apoptosis-related factors, inflammatory cytokines, and angiogenesis-related factors, although their effect on those targets can be very different. TTP is downregulated in several tumors, such as colon, kidney, pancreas, lung, stomach, amongst many others, acting as a tumor suppressor (Carrick & Blackshear, 2007). On the other hand, HuR acts as an oncogene and can be found up-regulated in various cancers, such as colon and pancreatic cancers (López De Silanes et al., 2003)(Abdelmohsen,

Srikantan, Kuwano, & Gorospe, 2008). Therefore, the recruitment of these cancer-related components to SGs can further contribute to the already established role of SGs in cancer.

1.8.3. Neurodegenerative diseases

SGs have been implicated in Alzheimer's disease (AD), Amyotrophic Lateral Sclerosis (ALS), Frontotemporal Dementia (FTD), Huntington's disease (HD), among other neurodegenerative diseases, namely by the existence of a link between SGs and the pathological hallmarks of neurodegeneration (Wolozin, 2012)(Wolozin, 2014). Mutations in SGs recruited-proteins can be found in patients with neurodegenerative disease and several studies demonstrated that SGs-related mechanisms are able to contribute to neuron loss by affecting neuronal functions. Other studies also showed that neuronal survival can be affected through the modulation of SGs dynamics by the pathological forms of SGs proteins (e.g. FUS, SMN, TDP-43) (Shukla & Parker, 2016)(Taylor, Jr., & Cleveland, 2016)(Maziuk, Ballance, & Wolozin, 2017). In line with these ideas, several proteins involved in the pathogenesis of neurodegenerative diseases are proteins recruited to SGs (Shukla & Parker, 2016)(Tara Vanderweyde et al., 2016), such as Tau, TDP-43, FUS, and Ataxin-2 (Atx-2), among others.

In axons, Tau stabilizes the microfilament network and promotes rapid axonal transport. But in pathological cases, like in the case of Alzheimer's disease, this protein forms toxic oligomers and fibrils (Kolarova, García-Sierra, Bartos, Ricny, & Ripova, 2012) and is able to accelerate SGs formation (Tara Vanderweyde et al., 2016). TDP-43, which a nucleating SGs component has several functions, such as alternative splicing in the nucleus and local translation. But alterations in the expression of this protein can cause neuronal loss and the development of diseases like FTD and ALS. Several studies have also shown that TDP-43 is able to affect the expression of proteins like Tau, Atx2, and FUS, implicated in both of these diseases (Sephton et al., 2011)(Polymenidou et al., 2011)(Tollervey et al., 2011)(Ferro, Yao, & Zarnescu, 2018)(Gu, Wu, et al., 2017)(Gu, Chen, et al., 2017). Also, the accumulation of pathological TDP-43 can be observed in AD, HD, and Dementia with Lewy Bodies (DLB) (Chen-plotkin, Lee, & Trojanowski, 2010). The increase of TDP-43 in the neurons' mitochondria can also contribute to the neuronal loss and to mitochondrial dysfunction (Wenzhang Wang et al., 2016). FUS can

bind to single- and double-stranded DNA, associated with RNA polymerase II (RNAP2) and III (RNAP3), engage in nucleo-cytoplasmic shuttling and is essential in cellular recovery (Lagier-Tourenne et al., 2012)(Ishigaki et al., 2012)(Tan, Riley, Coady, Bussemaker, & Manley, 2012)(Rogelj et al., 2012)(Zinszner, Sok, Immanuel, Yin, & Ron, 1997)(Morlando et al., 2012)(Wen-yuan Wang et al., 2013). Several studies have demonstrated that FUS overexpression and the presence of mutant FUS can lead to neuronal loss (Suzuki & Matsuoka, 2015)(Huang et al., 2011). Plus, studies with patients with FUS- positive inclusions in their neurons and glia have demonstrated a relationship between FUS mutations and ALS (Vance et al., 2013)(Vance et al., 2009)(Kwiatkowski et al., 2009).

But is important to notice that these FUS-related negative neuronal effects, happen not due to the presence of this protein, but instead due to the presence of its pathological form. In other words, neurodegeneration happens when the protein is overexpressed or when is mutated (Sharma et al., 2016).

When SGs are formed in response to stress, whereas defects in their dynamics can cause/enhance neurodegeneration. For example, in ALS patients, mutant TDP-43 can lead to hyper SGs assembly (Liu-Yesucevitz et al., 2010), whereas mutant Heterogeneous Nuclear Ribonucleoprotein A1 (hnRNPA1) and Heterogeneous Nuclear Ribonucleoprotein A2 (hnRNPA2) (both SGs components) can contribute to SGs persistence. Also, defects in the clearance mechanisms, like autophagy (or in the protein VCP) is another problem that contributes to a pathological persistence of SGs, which leads to neurotoxicity and neurodegeneration (Bentmann, Haass, & Dormann, 2013)(Monahan, Shewmaker, & Pandey, 2016).

Altogether, these studies suggest an important link between SGs and disease, hypothesizing about a possible implication of SGs in neurodegenerative diseases and therefore making them a possible target for disease treatment.

1.8.4 The RNA-binding proteins

RNA-binding proteins, also known as RBPs, are proteins able to bind to single or double-stranded RNA. These proteins are nuclear and cytoplasmic, but when in the nucleus they

tend to form ribonucleoprotein complexes, which are complexes of protein and pre-mRNA, also called hnRNPs (heterogeneous ribonucleoprotein particles) (Glisovic et al., 2008). These proteins have a great variety of functions in the cell, having a role in translation and in post-transcriptional control (Hogan et al., 2008) (Figure 1.4), being very important for RNA function, having a part in its biogenesis, transport, and stability (Glisovic et al., 2008). So far more than 1000 mammalian genes were identified as coding for RBPs, and 20% of all known proteins are RBPs (Gerstberger, Hafner, & Tuschl, 2014). More than 50% of known RBPs are expressed in the brain, where they are involved in different processes such as alternative splicing, transport, localization, and stability and translation of RNAs (Bryant & Yazdani, 2016). Besides being recruited to SG, several RBPs are components of neuronal RNA granules, also called transport ribonucleoprotein particles (RNPs), as they are motile granules transporting mRNA and containing several translational components (Kiebler & Bassell, 2006).

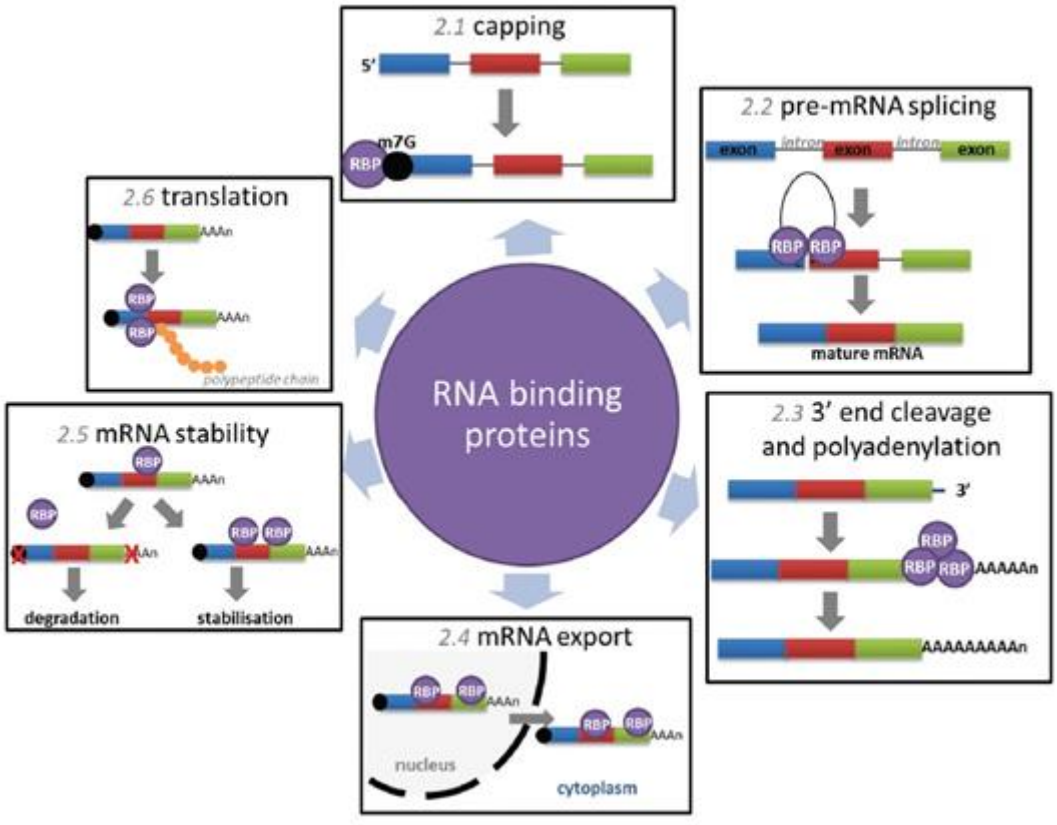


Figure 1.4 RNA-binding proteins (RBPs) regulate certain mechanisms of posttranscriptional control. RBPs can regulate capping by binding to the cap and promoting mRNA stability; pre-mRNA splicing, regulated by various RBPs within the macromolecular spliceosome; 3'-end cleavage and polyadenylation, where the addition of several adenosine residues is facilitated by a complex of RBPs; mRNA export, a shuttling that is mediated by the association of RBPs with specific transcripts; mRNA stability, modulated by the association between specific RBPs and certain transcripts; and translation, a process where RBPs are essential throughout. (Source: Sutherland, J. M., *et al*, 2015)

So, RBPs regulate several biological processes such as alternative splicing, a mechanism in which from the same gene different forms of mRNAs are produced and polyadenylation, a process that consists in the addition of adenylate residues in an RNA transcript and that is very important for transport and for translation efficiency. Also, RBPs have also an important role in RNA editing, like the protein ADAR, an RBP that catalyzes an enzymatic reaction which changes the nucleotides in the RNA sequence, consequently affecting mRNA transcripts (Glisovic et al., 2008). Furthermore, RBPs can be very important for the regulation of gene expression since they are involved in RNA localization. The mRNA localization ensures that transcription occurs in the intended site. For example, ZBP1 is an RBP that moves with B-actin mRNA to the cytoplasm and then localizes the mRNA to the region of the cell where it's supposed to be translated (Glisovic et al., 2008). RBPs are also able to control RNA transcripts due to the specific recognition of RNA targets (M.-H. Lee & Schedl, 2006). These proteins are able to control the generation and duration of the transcripts due to their RNA-binding domains, such as RNA-recognition motif (RRM), double-stranded RNA-binding motif (dsRBM) and zinc-finger motif (Stefl, Skrisovska, & Allain, 2005)(M.-H. Lee & Schedl, 2006). The RRM is a small protein domain composed of multiple structures that entail several types of interactions (RNA-RNA, protein-protein, and protein-RNA). This is the most abundant domain and has a role in a great variety of biological function, such as in post-transcriptional modification and RNA export and stability (Stefl et al., 2005). The dsRBM is a domain that was found to interact along the RNA duplex and that has specificity for several RNA structures due to its distinct chemical composition. This domain has essential roles in translational repression, RNA processing, RNA editing, etc (Stefl et al., 2005).

1.7.4.1 RBPs in disease

RBPs are proteins that not only constituted SGs but are also essential for their formation and for the cell homeostasis. As mentioned above, these proteins are involved in important cellular processes since many of them are able to interact with mRNA (Ravanidis, Kattan, & Doxakis, 2018). Therefore, the maintenance of physiological levels of RBPs is crucial for the correct cell function. For example, the proteins TIA-1/TIAR are very important for inflammation, cell cycle progression, apoptosis, among other

functions. Several studies have shown the existence of a relationship between TIA-1 and Tau (Apicco et al., 2018), where, in pathological conditions, TIA-1 promotes Tau aggregation and Tau affects TIA-1 RNA granule transport.

In neurons, RBPs have an essential role in gene expression and consequently in the brain complexity. These proteins and the factors with them associated are able to regulate processes such as splicing, transport, and translation in the neuron. So, it does not come as a surprise that when there is a disruption in the expression of these proteins, we can observe an association with the pathogenesis of neurodegenerative disorders such ALS, FTD, Fragile-X syndrome (FXS) and autism spectrum disorders (Ravanidis et al., 2018).

When it comes to cancer, a great number of RBPs have also been shown to be dysregulated in human cancers (Z. L. Wang et al., 2018). This difference in expression has been linked to aberrant alternative splicing (David & Manley, 2010)(Fredericks, Cygan, Brown, & Fairbrother, 2015)(Sebestyén et al., 2016), Copy Number Variations (CNV) like with IGF2BP2 in lung cancer, and protein mutations, like with U2AF1. (Imielinski et al., 2012)(Sebestyén et al., 2016)

1.8. Databases

In the 1960s, computers were faster and more capable. The availability of storage, like disks, increased, and with that so did the use of databases, contributing to an interactive sharing of information. (Date, 2004)

According to the Oxford English Dictionary, the first database dates back to the early 1960s. With the growth of technology, data became also electronically accessible. A database consists of a collection of related data, organized and stored electronically. Therefore, they serve as an important tool for research, as the number of databases has been increasing in the last years. The computer software is essential for the user to have access to the information in the database, more specifically the “database management system”, or DBMS. This software not only helps the organization of information but also contributes to the entry and storage of information. (Date, 2004)(Kroenke & Auer, 2012)

Throughout the years, the capability and the sizes of databases have grown in magnitude.

The progress in technology, in processors and computers, has allowed for an improvement in the performance of databases. (Korth & Silberschatz, 2011)(Kroenke & Auer, 2012)

Databases can have several types of content, such as text, statistics and multimedia and they can be classified based on their content, area or by technical aspect. For example, there are deductive databases, that combine logic programming with a relational database; active databases, that can respond to conditions inside and outside the database; cloud databases, that rely on cloud technology in which both the database and the software can be located remotely, amongst others (Korth & Silberschatz, 2011)(Kroenke & Auer, 2012).

CHAPTER 2 - OBJECTIVE

2. Objective of the work

Databases consist of storage of information that can be easily accessed and that is regularly managed and updated. Therefore, they serve as an important tool for research, as the number of databases has been increasing in the last years. There are several important databases on RBPs (K. B. Cook, Kazan, Zuberi, Morris, & Hughes, 2011)(Giudice, Sánchez-Cabo, Torroja, & Lara-Pezzi, 2016), focusing on different aspects of their structure or function, although they do not address the RBPs role/presence in SGs. Moreover, due to the growing interest in SGs research and their implication in human disease, there is an important unmet need to gather information about SGs and its components.

Therefore, the aim of this work was to create a storage of the collected data about SGs components. In other words, to catalog all proteins described so far as being part of stress granules in mammalian cells and for each protein collect information such as the complete protein name, if it is an RBP, molecular function, link to disease, cell type, and type of stimuli used induce SGs formation in the study where the protein was reported. Also, statistical information was obtained about topics such as molecular function, biological process, and protein class. For each protein there is available data on their expression profiles, where different groups ("disease" vs "control") were compared and analyzed using a dataset of human brain biopsy tissue sample, in the context of neurodegenerative diseases such as Alzheimer's, Parkinson's (PD) and Huntington's diseases and Amyotrophic Lateral Sclerosis. This extensive information collected was used to create an online database: The Mammalian Stress Granule Proteome (available at https://msgp.pt/)(Nunes et al., 2019), which is the first database to list and give information about all the stress granule proteins identified so far.

CHAPTER 3- METHODOLOGY

3. Methodology

3.1. Component identification and curation

To identify the SGs components described so far, an exhaustive search of all the published studies available on PubMed was done, using several keyword combinations, such as "Stress Granules AND sodium arsenite" or "Stress Granules AND mammals". All studies using mammalian cells were filtrated. The articles go from 1993 to 2017.

Each component was selected if the study demonstrated its recruitment to SGs through confocal microscopy or more recently also using co-immunoprecipitation with a core component of SGs. Each selected component was double checked. Together with the component information, the type of cells used the study, the stress stimulus used to induce the Sgs assembly and the reference of the study was also annotated.

All these information were gathered in the form of an Excel-based table.

3.2. SGs components details search

Aiming to provide a more complete description of the SGs component, which could be useful for the researchers using the database, several other informations were collected. For each protein, UniProt database was used as a source of information for the protein description, the protein name and its molecular function. UniProt (www.uniprot.org) is a database resource for protein sequence and function (Bateman et al., 2017). The information about SGs components Entrez ID and the chromosomal location was collected in the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov/gene). For some SGs components, the chromosomal location was also retrieved from Online Mendelian Inheritance in Man (OMIM) (www.omim.org), a catalog of all diseases with a genetic component. Information about the connection between protein and disease was also obtained through this database.

The classification of an SGs component as an RNA-binding protein was made according to the *Castello et al*., 2012 study (Castello et al., 2012) and the UniProt database. The *Castelo et al*. study defined the mRNA interactome in HeLa cells using protocols for covalent UV crosslinking of RBPs to RNA and therefore identified more than 800 RBPs. Furthermore, the RBPs identification was also confirmed using different online RBP

databases, the RBPDB and the ATtRACT database (K. B. Cook et al., 2011)(Giudice et al., 2016).

To identify which proteins of the MSGP Database were proteins implicated in autophagy, we used the information available in the Autophagy Database (Homma, Suzuki, & Sugawara, 2011), by searching for each one of SGs components in this database.

Like was already mentioned, we also obtained the information about the cell type and stress stimuli used in the first study describing the recruitment of that component to SGs.

The SGs component name and alias used in the database was retrieved using the GeneCards database (www.genecards.org), which has information about all known and predicted human genes.

To obtain information on the molecular and biological processes where SGs components are involved, the Protein Analysis Through Evolutionary Relationships (PANTHER) Classification System (http://www.pantherdb.org) was used, which is a database that classifies proteins to facilitate high-throughput analysis. Using the gene ID of each protein, Gene List Analysis for *Homo sapiens* was performed, to obtain the data with the molecular function, biological processes and protein class for all the components.

All these gathered details were organized in the form of Excel-based databases/tables.

3.3. Transcriptomic analysis

3.3.1. GEO DataSets search

The GEO DataSets database (www.ncbinlm.nih.gov/gds) was used to find the appropriate studies to perform a gene expression analysis for the identified SGs components. For that, the Advanced Search available in this database was used, searching for keywords such as "Alzheimer's disease AND gene profiling" or "Parkinson's disease AND gene profiling". To reduce the number of hits the search was narrowed down in the Organism setting, using the filter human.

A study for each of the four neurodegenerative diseases of interest (Alzheimer's disease – AD, Parkinson's disease – PD, Huntington's disease – HD, and Amyotrophic Lateral

Sclerosis – ALS) was chosen considering the number of samples and regions of the brain analyzed. So, for example, a study with more than 100 individuals and that analyzed more than one brain region would be preferable compared to a study with fewer samples and fewer brain regions studied, and for that chosen for the gene expression analysis.

For AD and HD, a study was found that although it only studied one brain region, the Dorsolateral prefrontal cortex (DLPFC), had a very high number of individuals sampled, more specifically, 624 samples (GSE33000). For PD, study with 114 individuals was found that analyzed different brain regions: medulla, striatum, frontal cortex and cerebellum (GSE28894). For ALS, the study used had 20 individuals and it analyzed the motor cortex brain region GSE4595.

3.3.2. Expression profiles/ GEO2R

The expression profiles for all stress granules components were extracted from a transcriptome dataset of human brain biopsy tissue sample, covering subjects with Alzheimer's disease, Parkinson's disease, Huntington's disease, Amyotrophic lateral sclerosis, and non-demented controls. Original expression data were analyzed using GEO2R web tool (http://www.ncbi.nlm.nih.gov/geo/geo2r/) comparing the different groups: AD versus Controls and HD versus Controls (GSE33000 (Narayanan M, Huynh JL, Wang K, Yang X et al. Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases. *Mol Syst Biol* 2014 Jul 30;10:743)), PD versus Controls (GSE28894), and ALS versus Controls (GSE4595 (Lederer CW, Torrisi A, Pantelidou M, Santama N et al. Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics* 2007 Jan 23;8:26)). An adjusted P-value of $P < 0,05$ accessed the SGs components whose expression was differentially expressed between both groups. Adjustments were made to correct the occurrence of false-positive results, using the Benjamini and Hochberg false discovery rate method.

For each SGs component, we retrieved the P-value, Adj P-values and Log fold-change (FC). Log FC gives us the ratio of the expression values ("disease" vs "control"), i.e., shows us the differential expression of each SGs component. Positive Log FC values

indicate up-regulation and negative values indicate down-regulation. These values were also included in the Excel tables.

Additionally, for each SGs component, the Log FC values were plotted into a graphic (GraphPad Software) to visualize the expression values of that component in each one of the studied neurodegenerative diseases.

3.4. Online database/catalogue

As mentioned, all the information collected was carefully organized into Excel sheets, with the columns for protein acronym, complete protein name, Gene ID, UniProt ID, chromosomal location, RBP, molecular function, OMIM disease, used cell type, used stimuli and study reference. This information was then used to build the open access online catalog: the Mammalian Stress Granule Proteome (MSGP) database (www.msgp.pt). This database was created using the Wordpress system, it was built from the Elementor Page Builder and using the tools Vue.js JavaScript framework, being completely integrated with the Contact Form 7.

3.5. Software tools and database implementation

The Mammalian Stress Granules Proteome (MSGP) is an online and open access database. The MSGP platform was conceived to work on all types of devices and integrates the functionality of user sign in/registration.

To implement the website the Wordpress system was used. The database was built using ‘custom fields’, based on the open source tool ‘Elementor Page Builder’.

Also, the database has features like a custom listing profile for each protein, custom fields with editing capability for each protein, highly customized GeneID cards, protein listing quick view, breadcrumbs navigation, a custom dashboard for front and end users and customized and multiple IDs for each protein. It was also included >50 widgets ready to use on the database, integrated in a clean system, compatible with PHP version 5.5+.

There were also used minified and combined assets to reduce the number of http requests and enhance load time and site performance.

The system was built using Vue.js Java Script framework and it was programmed clean and well-structured code, to facilitate access to the data.

CHAPTER 4- RESULTS

4. Results

4.1 SGs components and characterization

As mentioned, SGs are composed of stalled pre-initiation complexes: 40S subunits, translation initiation factors, poly(A)$^+$ mRNAs and RBPs. Nevertheless, SGs composition is different according to the type of stress and changes during the stress response (T Vanderweyde et al., 2013). Despite the existence of several studies reviewing SGs components, to date, there is none detailing all the components identified until now. Thus, we curated the literature for described SGs components covering the studies in mammalian cells. To the best of our knowledge, we annotated all the SGs components described so far, identifying 464 different proteins that are recruited to SGs (Annex B). This identification was made through manual curation of the found studies and the confirmation of the results showing that a specific component was recruited to SGs, either by confocal microscopy (co-localization with an SGs marker) or by co-immunoprecipitation with a core SGs component. To further characterize the SGs identified components, a gene ontology analysis of their molecular function was performed (Figure 4.1, Table 4.1).

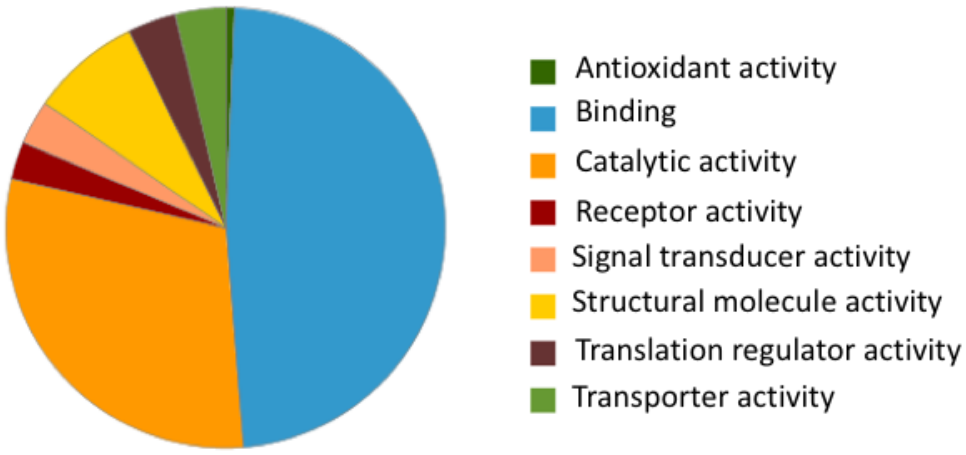


Figure 4. 1 Molecular functions of the 464 identified protein components of SGs, classified through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

RBPs are very important for SG assembly, due to their RNA-binding domains, which are essential for the aggregation and nucleation of SGs. More than half of the 464 identified proteins, more specifically 253 proteins (54%), were identified as RBPs, according to the Castello *et al.* study and the UniProt database. RBPs have low complexity domains, which

make them prone to aggregation and facilitates protein-protein interactions, possibly explaining this high prevalence of RBPs in SGs. When it comes to the molecular function, and by performing a molecular function analysis of the 464 SGs components (using the Protein Analysis Through Evolutionary Relationships available at http://www.pantherdb.org), the majority of the SGs components, more specifically 225 proteins, are binding proteins and 140 have catalytic activity.

Table 4.1 Number of SGs proteins according to the molecular function, obtained through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

Molecular Function	**Number of proteins**
Antioxidant activity	3
Binding	225
Catalytic activity	140
Receptor activity	13
Signal transducer activity	15
Structural molecule activity	38
Translation regulator activity	17
Transporter activity	17

Next, an analysis was performed to identify the biological processes in which the identified SGs components are involved. For that, the Gene List Analysis by the Panther Classification System was used, performing a biological processes analysis of the 464 SG components. From this analysis, two main processes stand out: the cellular and metabolic processes, with 308 and 235 SGs components, respectively. Interestingly, a high number of SGs components are also involved in the cellular component organization and in the biological regulation. A full description of the biological processes where SGs components are involved is described in figure 4.2 and table 4.2.

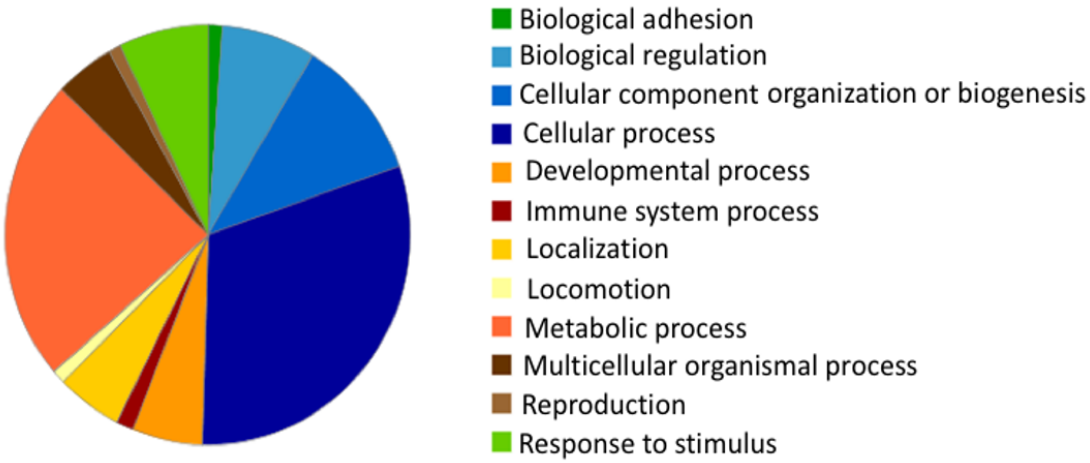


Figure 4.2 Biological processes where the 464 identified protein components of SGs are involved, classified through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

Table 4.2 Number of identified SGs proteins involved in different biological processes and obtained through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

Biological processes	**Number of proteins**
Biological adhesion	12
Biological regulation	75
Cellular component organization or biogenesis	110
Cellular process	308
Developmental process	56
Immune system process	14
Localization	52
Locomotion	11
Metabolic process	235
Multicellular organismal process	47
Reproduction	10
Response to stimulus	71

Finally, the protein class of the 464 identified SGs components was also analyzed. In line with the previous results showing that most of the SGs components are RBPs, 131 proteins belong to the class of nucleic acid binding proteins. However, the SGs components belong to a diverse range of proteins classes, including cytoskeletal proteins, hydrolases, enzyme modulators, transcription factors and transferases (Figure 4.3, Table 4.3).

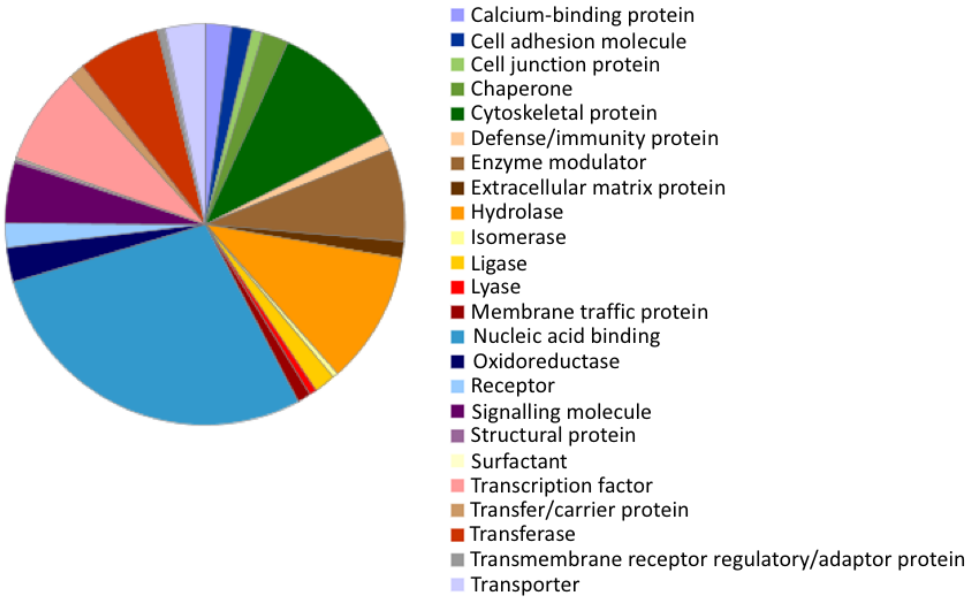


Figure 4.3 Protein class of the 464 identified SGs protein components, classified through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

4.2 SGs components and disease

As mentioned, RBPs have a wide range of functions being essential in the regulation of gene expression, post-transcriptional processes like pre-mRNA splicing, mRNA cytoplasmic export, turnover, storage, translation, and degradation. Due to these important functions, it is not surprising that deregulation in the expression of different RBPs seems to underlie a variety of human disorders, including cancer and neurodegenerative diseases (Lukong et al., 2008)(Cooper et al., 2009). In the last years, with the increase in transcriptomics analysis platforms, several studies were performed comparing gene expression levels across neurodegenerative diseases. They provided a new and valuable tool discovering pathways involved in the disease pathogenesis, and importantly identifying potential therapeutic targets and biomarkers for these diseases. In Amyotrophic Lateral Sclerosis (ALS), gene expression profiling has consistently implicated several cellular pathways in the disease pathogenesis, including cytoskeleton dysfunction, inflammation, and the impairment of transcription, of the ubiquitin-proteosome system and of the mitochondria (Cooper-Knock et al., 2012). In Parkinson's disease (PD) or Alzheimer's disease (AD), gene expression profiling also showed several alterations in cellular pathways with important consequences for the disease pathogenesis (Cooper-Knock et al., 2012).

Table 4.3 Number of SGs proteins in each protein class, obtained through Gene List Analysis by the PANTHER Classification System (http://www.pantherdb.org).

Protein Class	Number of proteins
Calcium-binding protein	10
Cell adhesion molecule	8
Cell junction protein	4
Chaperone	10
Cytoskeletal protein	50
Defense/immunity protein	6
Enzyme modulator	35
Extracellular matrix protein	6
Hydrolase	50
Isomerase	2
Ligase	8
Lyase	3
Membrane traffic protein	5
Nucleic acid binding	131
Oxidoreductase	13
Receptor	9
Signaling molecule	23
Structural protein	1
Surfactant	1
Transcription factor	36
Transfer/carrier protein	6
Transferase	31
Transmembrane receptor regulatory/adaptor protein	3
Transporter	15

Given the important role of RBPs in cellular functions, it is expected that deregulation in their expression could have a profound effect on neuronal health, contributing to the deregulation of different pathways underlying neurodegenerative diseases pathogenesis. In view of these ideas one important question should be outlined: is the expression profile of SG components altered in the context of different neurodegenerative diseases?

Trying to answer this question we investigated published data from human brain samples, comparing the gene expression profiles of the identified SGs components in for neurodegenerative diseases with data from non-demented controls, using Geo2R web tool (available at http://www.ncbi.nlm.nih.gov/geo/geo2r/). The expression profiles for all SGs components were extracted from a dataset of human brain biopsy tissue sample, with subjects with AD, PD, HD, ALS, and non-demented controls. This data was analyzed and the “disease vs controls” groups (AD versus Controls, HD versus Controls, PD versus

Controls and ALS versus Controls) were compared. The transcriptomic analysis showed us which proteins were differentially expressed in the neurodegenerative diseases studied. The data obtained was plotted into a graphic and is also a part of the component-specific page.

The proteins that are overexpressed/repressed in one disease could not necessarily be overexpressed/repressed in another disease too. For example, Ago1 is induced in AD but repressed in HD, while TIA-1 is repressed in AD and induced in HD (Figure 4.4).

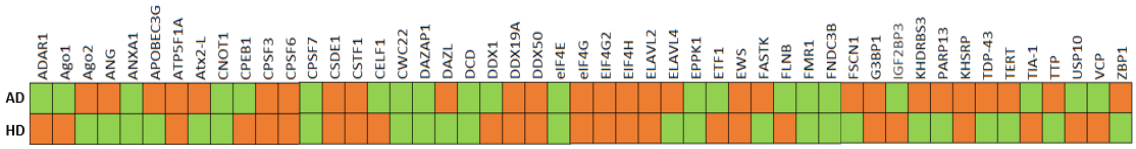


Figure 4.4 Heat-map of the differential expression in 50 selected SGs components. Example of 50 SGs RBPs that are Induced ($p<0.05$) (green) / Repressed ($p<0.05$) (orange) in Alzheimer's disease (AD) and Huntington's disease (HD). Through the GSE33000 study, we obtained the p-values and Log FC values for all SGs proteins. The proteins that had a significant p-value ($p<0.05$) were selected.

As previously mentioned, SGs were implicated in different neurodegenerative diseases by several studies. Plus, the majority of SG components are RBPs, proteins that are very important for posttranscriptional control and translation of RNAs in neurons (Bryant & Yazdani, 2016). Therefore, alterations in their expression may have an impact on the pathogenesis of neurodegenerative disorders. According to our analysis, 380 proteins were found to be differentially expressed in AD. From those, a total of 187 were induced and 193 were repressed. In HD, 395 proteins were found to be differentially expressed, being 195 of those proteins induced and 200 repressed. There were 191 proteins that we found to be induced/repressed in both diseases. From those, 90 proteins were found to have the expression increased and 101 proteins were repressed in both AD and HD (Figure 4.5).

The real importance of these results in terms of disease pathogenesis is not clear, although this major dysregulation in the SGs components expression might be relevant for these diseases. For example, in line with this idea, a recent paper by Johnson and colleagues (2018) show a strong impairment in RBPs expression in the brain of Alzheimer's disease patients, which is associated with the disease (Johnson et al., 2018).

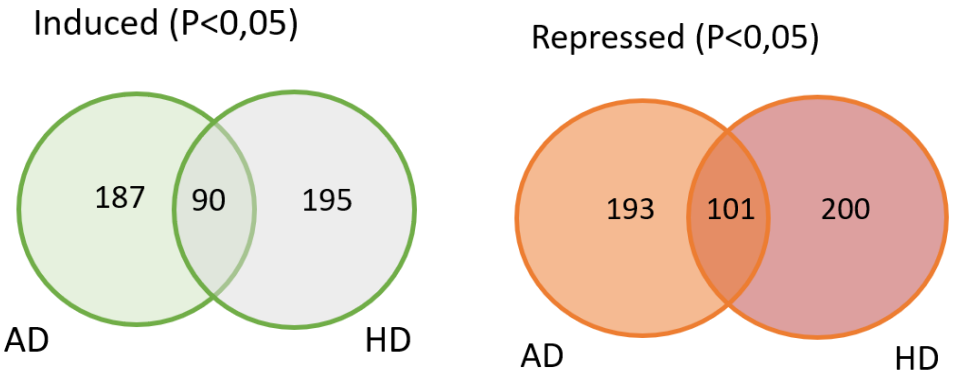


Figure 4.5 Differential expression of SGs components in Alzheimer's and Huntington's disease. Number of proteins differentially expressed in Alzheimer's disease (AD) and Huntington's disease (HD), according to the transcriptomic analysis. Through the GSE33000 study, we obtained the p-values and Log FC values for all SGs proteins. The proteins that had a significant p-value ($p<0.05$) were selected.

4.3 Other details about SGs

SGs components are variable, depending on the cell type and stress stimuli used to induce their formation. There is a wide range of different cell types and different stressors to induce SGs assembly reported. In the studies, we selected for the SGs components identification, the main cell type that was used to promote SGs assembly was the U-2OS cells. From all the identified components, 327 proteins were reported to be recruited to SGs using these cells, whereas 121 proteins were studied in HeLa cells. It is important to note that some of the proteins were studied in both (Figure 4.6). Also, 78 more different types of cells were used other than U-2OS and HeLa cells. The complete list of all cell types can be found in Annex C.

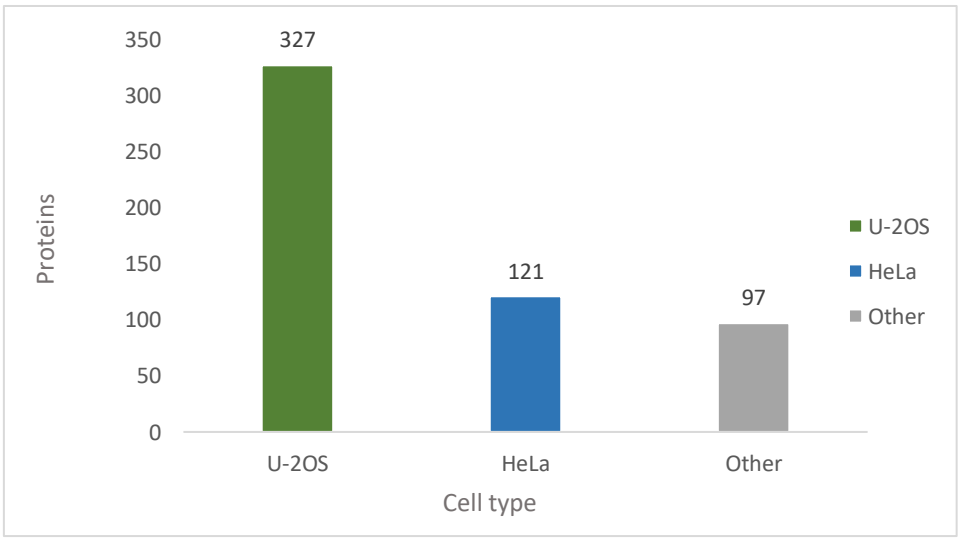


Figure 4.6 Number of identified SGs proteins according to the cell type used in the study (U-2OS and HeLa cells, and other cell types).

Another important issue in SGs study is the stimulus used to induce cellular stress. From literature search, we found that the main stress stimuli used to induce SGs assembly in the identified studies was sodium arsenite: 427 studies used this stressor to induce SGs formation, reflecting that the same number of proteins was shown to be recruited to SGs using that stressor. From all the SGs components, we found that 63 proteins were studied through SGs induction by heat shock and 66 by other 26 different types of stressors (Figure 4.7), such as Hippuristanol, Dithiothreitol, and Thapsigargin. The complete list of the used stressors can be found in Annex D.

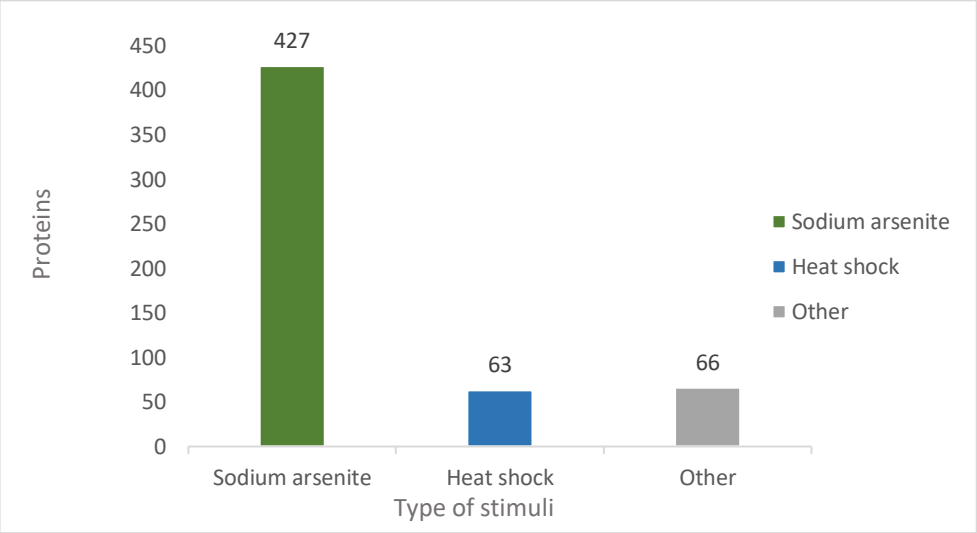


Figure 4.7 Number of proteins identified using different stress stimulus to induce SGs assembly (Sodium arsenite, heat shock, and other stressors).

Finally, we also gathered l the information on each SGs component chromosomal location (gene). There a distribution of the genes across all chromosomes, being, however, more found in chromosome 3, whilst only 2 proteins are localized in chromosome 21 (Table 4.4).

Table 4.4 Number of proteins which genes are expressed in each chromosome, based on the information obtained through the NCBI database and OMIM database.

Chromosome	Number of proteins
X	10
1	19
2	17
3	23
4	8
5	16
6	6
7	16
8	17
9	16
10	16
11	14
12	21
13	9
14	12
15	11
16	9
17	19
18	3
19	14
20	13
21	2
22	14

4.4. Database implementation

4.4.1. Database structure

To make available the huge amount of important data collected about SGs, an online open-access database was created, at https://msgp.pt/. The MSGP database structure was planned to work in a very intuitive and user-friendly manner. The ‘How to use’ section allows users to find a tutorial of the search process and the system navigation through the database and in each SG component-specific page. Also, all SG components can be found in the ‘Protein Index’ page, where they are listed alphabetically.

The MSGP database contains the 464 proteins identified as components of the SGs, which can be retrieved through different pages from the database or from a general search. The structure of the database is described in Figure 4.8, providing two major sections: all proteins, grouping the information and details on the 464 proteins identified, the RBP section, where all SGs components classified as RBPs are grouped, and the Autophagy

section, where all the autophagy-related proteins are grouped. The database allows further exploring of the identified and curated components, in the 'Featured proteins' section or in the section where more recent proteins were added to the database ('New proteins' section). Furthermore, the database allows the exploring of listed proteins according to different tags, such as sodium arsenite, U-2 OS or HeLa cells.

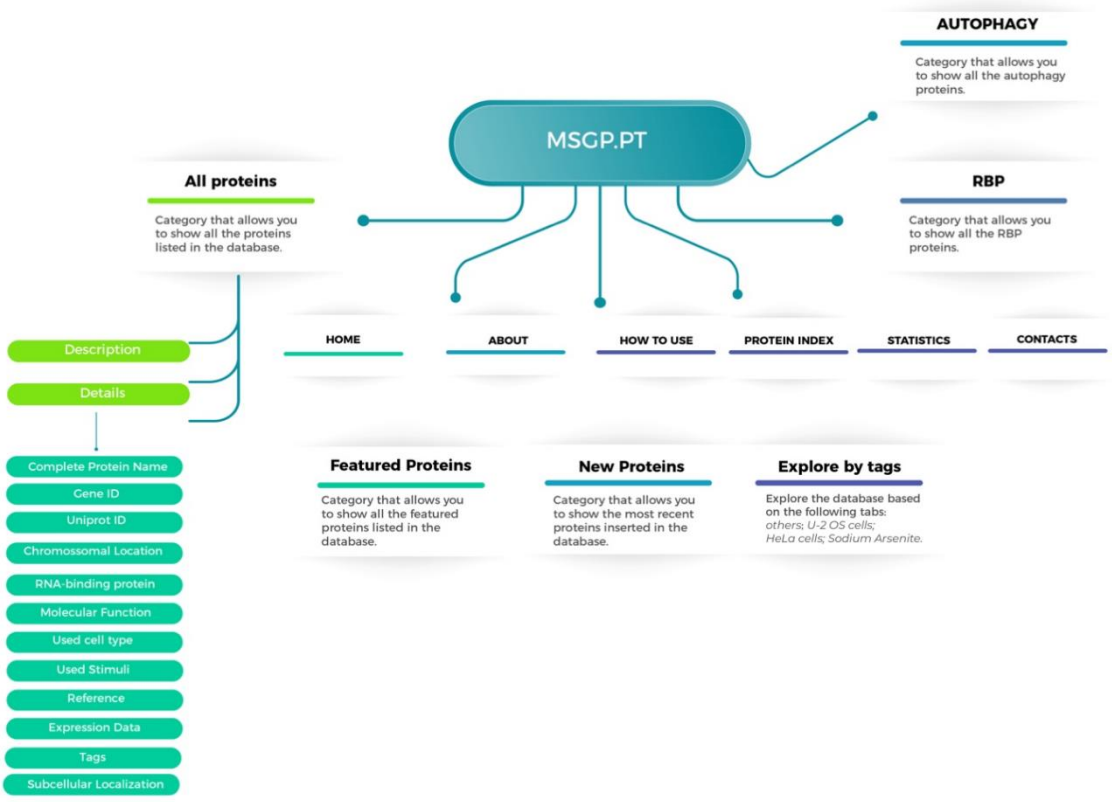


Figure 4.8 The diagram illustrates the structure of the database. The white boxes indicate the categories (All proteins, RBP, Autophagy) and features (Home, about, how to use, protein index statistics, contacts, featured proteins, new proteins, explore by tags) available on the home page. The green boxes represent the information available in the protein profiles.

4.4.2 Database content and annotation

The database comprises detailed information about the identified components of SGs in mammalian cells. For each protein, information was obtained from public databases: i) a small description of each protein; ii) identification details, such as protein name, gene ID, UniProt ID, chromosomal location, if it is an RNA-binding protein (RBP), if it is an

autophagy-related protein, molecular function, OMIM disease, cell type used in the study, stress stimulus used and the reference of the original study describing the recruitment of the component to SGs. Moreover, for each SGs component, the expression levels in the context of different neurodegenerative diseases were also gathered. This detailed information was retrieved from public databases. As an example, the gene ID was obtained in the National Center for Biotechnology Information (NCBI) database (www.ncbi.nlm.nih.gov/gene), and the chromosomal location was retrieved from the NCBI database and *Online Mendelian Inheritance in Man* (OMIM) (www.omim.org). The protein classification as an RBP was obtained through the study developed by Castello, and colleagues (Castello *et al.*, 2012) and from the UniProt database (www.uniprot.org) and confirmed in RBP databases (RBPDB and ATtRACT databases). The cell type and stress stimuli information in which the SGs component was identified were retrieved from the original study describing the recruitment of that protein to the SGs.

4.4.3 Online interface and query of the database

The online interface to the database is organized with a "Search" bar and four categories. To find all the proteins in this database the user can scroll through the database or go to "All proteins" in the home page. But, to look for one specific protein, the user must type the name of the protein in the search bar and just click on "Search" (Figure 4.9). The four categories are entitled: "All proteins", "RBP", "OMIM disease" and "Autophagy" (Figure 4.10). The categories are set to present and group the search results, allowing the user to narrow the search. The user can find all the proteins in the database by selecting the "All proteins" category, while all the RNA-binding proteins can be found by selecting the "RBP" category. The "OMIM disease" category gives all the proteins associated with pathologies and all autophagy-related proteins are grouped in the "Autophagy" category.

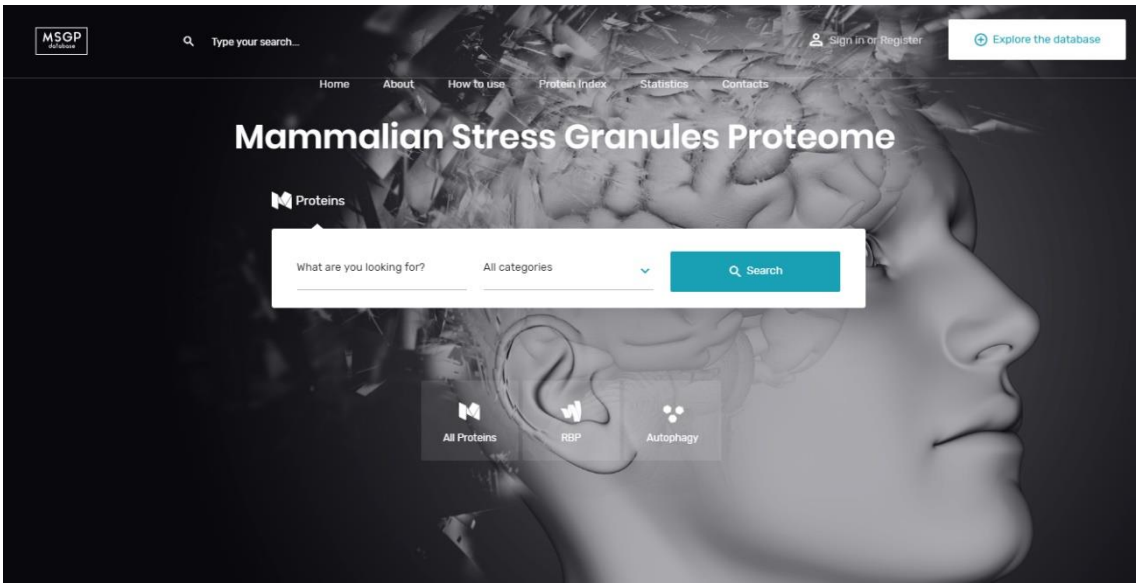


Figure 4.9 Home-page of the MSGP database. The landing page of the MSGP database, with the search form and the different pages.

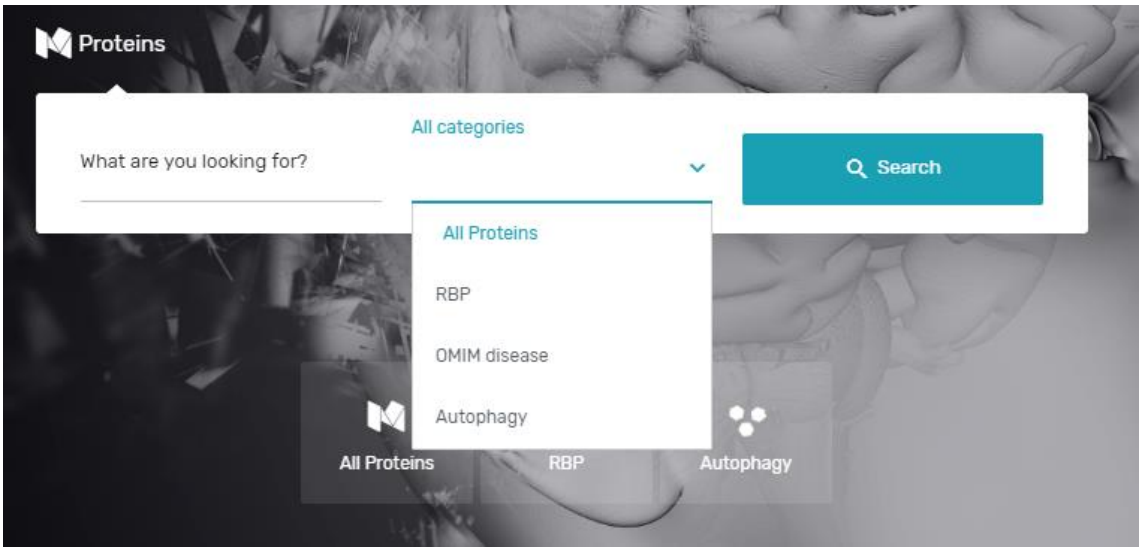


Figure 4.10 All categories of the MSGP database. The different options to narrow the search.

Furthermore, the proteins are also organized with tags. By going to the end of the home page, the user can explore the database by the use of tags, such as HeLa Cells or Sodium arsenite (Figure 4.11). Also, by clicking on "All proteins" in the home page, the search can be narrowed down with the use of filters (right top of the page).

More information about the database, how to use and statistics about the SGs proteins, can be found in the "About", "How to use" and "Statistics" sections, respectively.

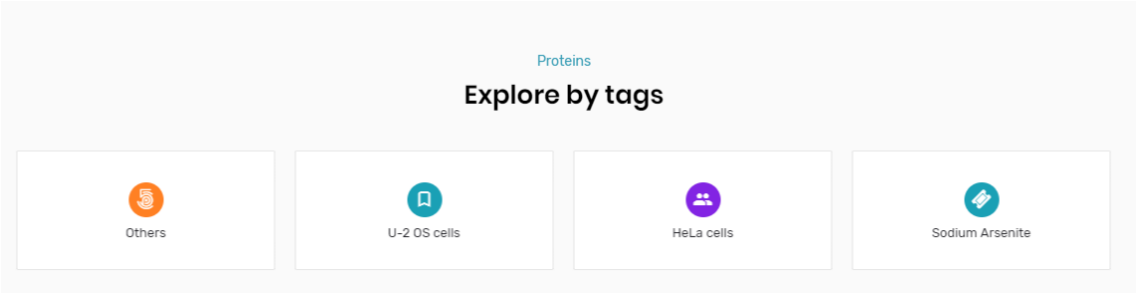


Figure 4.11 Explore by tags. The possibility of accessing the listed components through different tags.

4.4.4. Perform a search and search results

To search the database, several types of filters can be used: RBP, disease, cell type used (such as HeLa cells or U-2OS cells) and stress stimuli used (such as sodium arsenite or heat). The search will result in a list of proteins profiles.

To search for a specific protein, the user should use the abbreviation of the protein name. For example, to find the Fused in Sarcoma protein, the user should type FUS in the search bar. If no category selection is made, the search is performed in the 'All Proteins' category. Is important to reference that often a protein is known by other different names (aliases), so if the protein the user is looking for does not appear, it could mean that the protein is under a different alias. Aliases can be found in databases such as GeneCards (https://www.genecards.org/) or UniProt (https://www.uniprot.org/).

Only the component must be typed in the search bar, since the use of additional words, such as "and" or "or" will not be recognized by the database. In the case of certain components being listed under similar names, like is the case of the components of the PABP family, for example, by typing "PABP" in the search bar, all those components will appear. So, in this example, the user will have three hits: PABP1, PABPC3, and PABP4.

The profiles of each SGs protein that appear as a result of the search can be visualized by clicking on it. In each page, a brief description and details on the protein, the categories, and tags of the protein, and the results of the transcriptomic analysis are available (Figure 4.12).

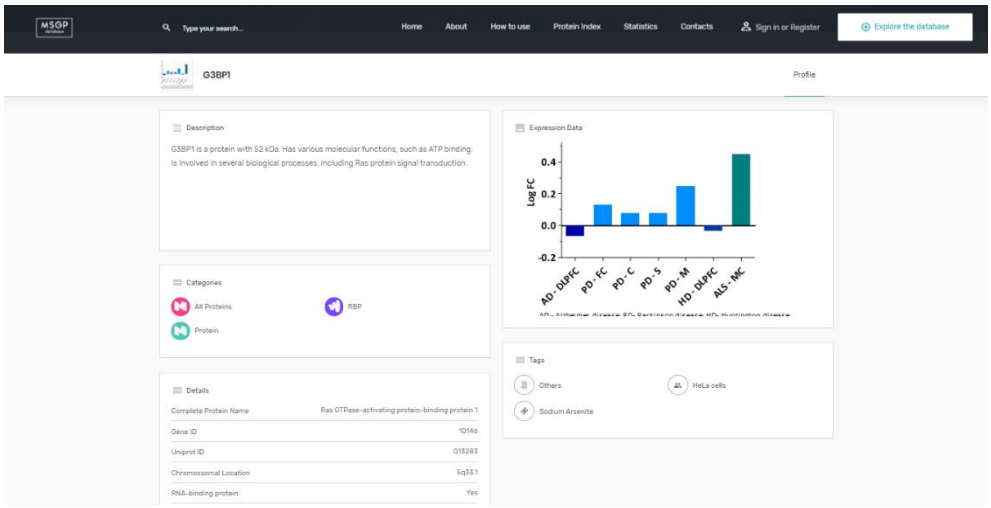


Figure 4.12 The profile of one protein, component of SGs (G3PB1) in the MSGP database.

4.4.5 SGs components details page

As already mentioned, information was gathered for each SGs component, which are grouped together in the "Details" section in their specific page and includes the complete protein name, Gene ID, UniProt ID, chromosomal location, RBP classification (Yes or No), the molecular function, OMIM details, the reference to the original study, the cell type and type of stress stimuli used (Figure 4.13).

Details

Complete Protein Name	T-cell intracellular antigen-1
Gene ID	7072
Uniprot ID	P31483
Chromossomal Location	2p13.3
RNA-binding protein	Yes
Molecular Function	AU-rich element binding ,poly(A) binding ,RNA binding
OMIM Disease	Welander distal myopathy
Used cell type	DU145, COS-7, HeLa
Used Stimuli	Sodium arsenite, heat shock, sorbitol, UV irradiation

Reference
Kedersha, N. L.; Gupta, M.; Li, W.; Miller, I.; Anderson, P. (1999). RNA-Binding Proteins Tia-1 and Tiar Link the Phosphorylation of Eif-2α to the Assembly of Mammalian Stress Granules. The Journal of Cell Biology, 147(7), 1431–1442

Figure 4.13 Details from the SGs component TIA-1, as presented in the MSGP database.

Furthermore, in each component-specific page, we can find a graphic with the expression profile for the SG component. For example, by accessing the page of G3BP1 we can observe that G3BP1 is overexpressed in PD and ALS and repressed in AD and HD (Figure 4.14), whereas TTP is overexpressed in 3 brain regions in PD, in HD and ALS and repressed in AD and on the medulla in PD (Figure 4.15).

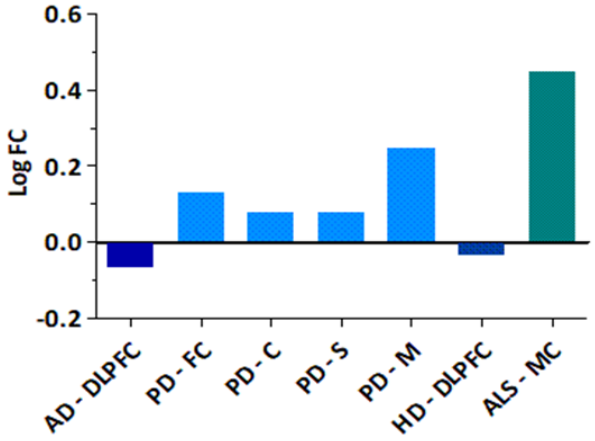


Figure 4.14 Expression profile of G3BP1 in different neurodegenerative diseases (AD, PD, HD, and ALS) in different brain regions. Through the studies GSE33000, GSE28894, and GSE4595 we obtained the p-values and Log FC values for all SGs proteins. The Log FC values were used to build graphics of expression for each protein. AD- Alzheimer's disease; PD- Parkinson's disease; HD- Huntington's disease; ALS- Amyotrophic lateral sclerosis; DLPFC- Dorsolateral prefrontal cortex; FC- Frontal cortex; C- Cerebellum; S- Striatum; M- Medulla; MC- Motor Cortex

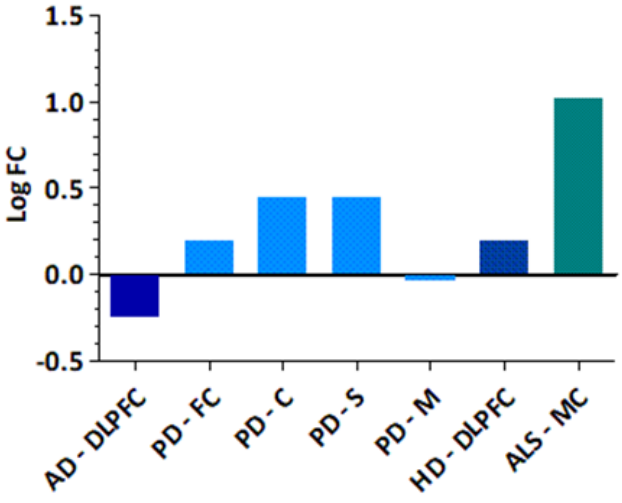


Figure 4.15 Expression profile of TTP in several neurodegenerative diseases (AD, PD, HD, and ALS) in different brain regions. Through the studies GSE33000, GSE28894, and GSE4595 we obtained the p-values and Log FC values for all SGs proteins. The Log FC values were used to build graphics of expression for each protein. AD- Alzheimer's disease; PD- Parkinson's disease; HD- Huntington's disease; ALS- Amyotrophic lateral sclerosis; DLPFC- Dorsolateral prefrontal cortex; FC- Frontal cortex; C- Cerebellum; S- Striatum; M- Medulla; MC- Motor Cortex

4.4.7. User interaction and future updates

Although the database does not have a page for users to submit new SG components or information relating the components, researchers can send that information to the contacts in the database. Also, the entire data set of the database is available for any researcher upon request by email.

In terms of future updates, information on the gene expression levels of SG components for different types of cancer will be added, together with a continuous update for new SG components from the published studies indexed in PubMed. Also, we consider including other SG components besides proteins, such as mRNAs and miRNAs.

Furthermore, each protein details will continue to be updated with new and relevant information, such as the number of studies where that particular component was described.

CHAPTER 5- DISCUSSION

5. Discussion

5.1 SGs relevance

SGs are formed during stress responses, such as oxidative stress, heat shock, and nutrient deprivation. These stress stimuli are capable of inhibiting translation initiation and SGs contribute to the improvement of cell viability during the stress response (Protter & Parker, 2016). Although SGs were once considered nonspecific aggregates with limited importance in the cellular response to stress, nowadays their importance in human disease is clear and they have been associated to cancer, neurodegenerative diseases, viral infections, immune diseases and aging (Anderson & Kedersha, 2008). SGs were discovered approximately 30 years ago, and since then a lot of progress has been made in this field, given us information about the proteins and RNAs that composed them, the interactions that lead to their assembly or alterations that lead to human disease and how those alterations and the SG components connect (Treeck & Parker, 2019). The SGs research has been growing in interest each year, as in the last 5 years more than 900 articles about stress granules were published. These are the most studied RNP granules, mainly due to their biological importance and their connection to disease. For that, the construction of an online database curating and centralizing all the information about the SGs is important and timely. In fact, there is a growing trend in defining new tools and establishing guidelines for the study of SGs. For example, Van Treeck and Parker (Treeck & Parker, 2019) recently reviewed and critically discuss the different imaging methods to study SGs. Therefore, the development of new tools that are freely available to researchers worldwide also potentiates SGs research, allowing a more profound understanding of its biology and implication for human disease.

5.2 SGs components

SGs are constituted by several components, including different types of RNAs, proteins and signaling molecules. In this work, we focused on proteins, which are the components for which there is more information available. But the composition of this granules can vary depending on the cell type and stress stimuli used to induce SGs formation (M. G. Thomas et al., 2011)(J. R. Buchan et al., 2011). For that reason, it was also important to have this information detailed for each protein.

From the different SGs components, identified RBPs stand out as the most common proteins recruited to these cellular foci. RBPs are implicated in an abundance of RNA processing events and throughout the years has become clear that these proteins play a critical role in the neuronal function and homeostasis. These proteins have been related to functions like RNA editing, nucleo-cytoplasmic trafficking, alternative splicing, and miRNA biogenesis. Moreover, several studies connected the deregulation of their levels to neuronal dysfunction and neurodegeneration (Ravanidis et al., 2018). RBPs have RNA-binding domains, which are essential for SG formation. These proteins are very important for protein aggregation and consequently for SG nucleation, being able to facilitate the recruitment of other proteins to SGs (Gilks et al., 2004)(Bounedjah et al., 2014). This aligns with the fact that through our analysis of the SG components' molecular function, it was possible to see that a majority of SG components is a binding protein, which is very important for SG assembly due to their aggregation domains. Plus, it also aligns with the fact that 131 proteins belong to the class of nucleic acid binding proteins.

Given the important role that these proteins have in gene expression and neuronal homeostasis and the fact that several studies hypothesized about their association with the pathogenesis of neurodegenerative diseases (Alves et al., 2016)(Johnson et al., 2018), it becomes important to better understand their interactions with RNA and how the recruitment of these proteins to SGs are important for cell function.

5.3 On the SGs studies

The great majority of SG components were reporting to being recruited to these foci through oxidative stress, more specifically, using sodium arsenite. Sodium arsenite was the first known stressor used to induce SGs assembly and it is the most used not just due to its ease of use, but also because it is considered the most effective agent to induce SG assembly (Kedersha et al., 1999).

The main cell type used to study SG assembly in the identified studies was U-2OS cells. This cell line is known for its fast growth ability and its high transfection efficiencies (Lauvrak, Munthe, Kresse, Stratford, & Namløs, 2013). But although the majority of proteins was identified through the use of U-2OS, there was a greater variety of studies using HeLa cells. These cells are famous for their immortality, being able to proliferate

indefinitely. They are well studied, fast-growing, having a rapid multiplication rate, the culture handling is technically easy, is a readily available and cost-effective cell line. (Rahbari, Sheahan, Modes, Collier, & Macfarlane, 2009)(Capes-davis et al., 2010) However, due to the fact that SGs are membraneless compartments, the use of this cells to study their dynamic can be a challenge, which affects the identification of effective therapeutics (Marrone et al., 2018). Fortunately, studies using patient cells and iPSC derived cell lines are starting to appear, which hopefully will lead to a future where we may have a profile of SG components according on the stimulus and/or cell type (Lenzi et al., 2015)(Marrone et al., 2018)(Zhang et al., 2019).

All the identified studies used confocal microscopy and/or co-immunoprecipitation to identify the SG proteins that were recruited to SGs. Confocal microscopy allows the localization and identification of the SGs components in the cell (Paddock, 1999), while co-immunoprecipitation is a technique generally used to analyze protein-protein interactions and allows the identification of a protein with the use of target protein-specific antibodies (Phizicky & Fields, 1995). These techniques are considered the gold standard to study the components recruitment to the SGs.

5.4 SGs and disease

With each passing year, there is more and more evidence about the fact that SGs have indeed a part to play in various diseases, such as neurodegenerative diseases. Several studies, such as Wolozin, B. et al, 2012, have pointed to the fact that several of the SGs proteins are involved in the pathology of diseases like Alzheimer's disease, Parkinson's disease and Amyotrophic Lateral Sclerosis, and that the pathological presence of this granules have shown to be involved in neurodegeneration and neurotoxicity, making SGs promising therapeutic targets (Wolozin, 2014). Even late last year, another study related the SGs dynamics with spinocerebellar ataxia type 2 (Paul, Dansithong, Figueroa, Scoles, & Pulst, 2018), showed its possible relevance in other neurodegenerative diseases.

Through our transcriptomic analysis, where the 'disease' and the 'control' groups were compared, we were able to see how many and which proteins were differentially expressed in the four neurodegenerative diseases. The overexpression and underexpression of certain proteins are important given that these alterations in protein

expression can have a relevant role in neurodegenerative disease. For instance, it has been previously reported that certain changes in protein expression characterize the pathophysiology of AD. Some proteins identified as risk factors for AD were found to be differentially expressed in AD phenotypes. (Johnson et al., 2018) Further analysis of the protein expression in this disease will lead to more insights in the cellular and biochemical changes that happen in the brain during AD, which for now, still remains poorly understood.

SGs are a process of regulated protein aggregation, which contrasts with the pathological and dysregulated aggregation observed in several neurodegenerative diseases, such as the Polyglutamine diseases (Matos, de Almeida, and Nóbrega, 2017). Several studies provide evidence for this possible implication of SG and their components in polyQ pathogenesis. For example, it was shown that Atx-2 is able to interact directly with mRNAs (Yokoshi et al., 2014), and directly or indirectly with several proteins (Ralser, Albrecht, et al., 2005)(Ralser, Nonhoff, et al., 2005). Moreover, mutations in several SG components increase their propensity to aggregate and to induce SG formation (Wolozin, 2012). The possible link between pathological protein aggregates and SG is further supported by evidence of co-localization between them. In a model of brain ischemic injury, one day after the insult ubiquitin-containing aggregates were detected in neurons, but they did not co-localize with SG. Interestingly, at 2 days after the injury, it was observed a co-localization between protein aggregates and SG (DeGracia, Rudolph, Roberts, Rafols, & Wang, 2007). Moreover, SG components are found in the neuropathological protein aggregates of polyQ diseases (Waelter et al., 2001)(Furukawa, Kaneko, Matsumoto, Kurosawa, & Nukina, 2009)(Elden et al., 2010), and in other neurodegenerative diseases (Dewey et al., 2012). For that, the study of SGs offers also a model and an opportunity to study the aggregation process in these incurable diseases.

5.5 The MSGP database

Databases consist of storage of information that can be easily accessed and that is regularly managed and updated (Curbelo, Loza, De Yébenes, & Carmona, 2014). With the growth and accumulation of biological data, mainly due to higher-throughput and lower-cost DNA sequencing technologies, came the need to create a vast number of databases, at a fast rate, with the purpose to support in human-related research (Zou, Ma,

Yu, & Zhang, 2015). Therefore, they serve as an important tool for research, as the number of databases has been increasing in the last years. In 2014, there were 1552 databases publicly accessible online (Fernández-Suárez, Rigden, & Galperin, 2014). Importantly, databases aim not just to store and organize data, but also to share data in a searchable manner, facilitating data retrieval for researchers (Zou et al., 2015). Databases are an easy and cheap way to organize great volumes of data, can store enormous amounts of data and also allow computers to integrate and exchange data in an automated manner is why nowadays, for which they are crucial and more and more indispensable for researchers (Zou et al., 2015)(C. E. Cook et al., 2016).

The MSGP database is new and unique, being the first database to provide organized information about stress granules components, such as their molecular function if the protein is an RBP and their link to disease. Importantly, it also provides the expression data of those stress granules' components in the context of different neurodegenerative diseases, showing which proteins are overexpressed and underexpressed in different brain regions for AD, PD, HD, and ALS. Since the publication of the paper describing the database we already received important feedback from the researchers in the area, acknowledging the importance of the database for the SGs research field or even proposing future updates.

CHAPTER 6 - CONCLUSION AND FUTURE PERSPECTIVES

6. Conclusion and future perspectives

In the last years, there has been a growing interest in SGs research, due mainly to their implication in different human health conditions. But although we are seeing an increase in reviews on the topic, there is still a lack of resources and tools for their study.

For the first time, there is a catalogue that centralizes and storage all the information about the stress granules' components. This huge effort yielded a unique database, which will provide a valuable tool for researchers in the area.

Furthermore, it has the possibility to be updated, with not just new information about the existing proteins, such as gene expression profiles in other diseases, like for example cancer. It will also continuously updated with new components described has being recruited to SGs. Moreover, it has also the possibility of including non-proteins components such as microRNAs or mRNAs, which are also present in the SGS.

In the future, with the continuous profiling of SGs components in patient-derived cells, it will be very important to establish composition profiles for SGs according to the type of cell, disease or stress stimulus. This would be an important resource for understanding the SGs biology and if possible will be also implemented in the database.

Therefore, there are plans for continuous improvement and update of the MSGP database, which will constitute an important asset for the SGs research field.

REFERENCES

Abdelmohsen, K., Srikantan, S., Kuwano, Y., & Gorospe, M. (2008). miR-519 reduces cell proliferation by lowering RNA-binding protein HuR levels. *Proceedings of the National Academy of Sciences*, *105*(51), 20297–20302. https://doi.org/10.1073/pnas.0809376106

Ackerman, D., & Simon, M. C. (2014). Hypoxia, lipids, and cancer: surviving the harsh tumor microenvironment. *Trends Cell Biol.*, *24*(8), 472–478. https://doi.org/10.1016/j.pain.2013.06.005.Re-Thinking

Adams, D. R., Ron, D., & Kiely, P. A. (2011). RACK1, A multifaceted scaffolding protein: Structure and function. *Cell Communication and Signaling*, *9*(1), 22. https://doi.org/10.1186/1478-811X-9-22

Adjibade, P., St-Sauveur, V. G., Huberdeau, M. Q., Fournier, M., Savard, A., Coudert, L., … Mazroui, R. (2015). Sorafenib, a multikinase inhibitor, induces formation of stress granules in hepatocarcinoma cells. *Oncotarget*, *6*(41), 43927–43943. https://doi.org/10.18632/oncotarget.5980

Aitken, C. E., & Lorsch, J. R. (2012). A mechanistic overview of translation initiation in eukaryotes. *Nature Structural and Molecular Biology*, *19*(6), 568–576. https://doi.org/10.1038/nsmb.2303

Alves, S., Biferi, M.-G., Cartier, N., El Hachimi, K., Furling, D., Ruberg, M., … Sittler, A. (2016). Lentiviral vector-mediated overexpression of mutant ataxin-7 recapitulates SCA7 pathology and promotes accumulation of the FUS/TLS and MBNL1 RNA-binding proteins. *Molecular Neurodegeneration*, *11*(1), 1–20. https://doi.org/10.1186/s13024-016-0123-2

Anderson, P., & Kedersha, N. (2002a). Stressful initiations. *Journal of Cell Science*, *115*(Pt 16), 3227–3234.

Anderson, P., & Kedersha, N. (2002b). Visibly stressed: The role of eIF2, TIA-1, and stress granules in protein translation. *Cell Stress and Chaperones*, *7*(2), 213–221. https://doi.org/10.1379/1466-1268(2002)007<0213:VSTROE>2.0.CO;2

Anderson, P., & Kedersha, N. (2006). RNA granules. *Journal of Cell Biology*, *172*(6), 803–808. https://doi.org/10.1083/jcb.200512082

Anderson, P., & Kedersha, N. (2008). Stress granules: the Tao of RNA triage. *Trends in Biochemical Sciences*, *33*(3), 141–150. https://doi.org/10.1016/j.tibs.2007.12.003

Anderson, P., Kedersha, N., & Ivanov, P. (2015). Stress granules, P-bodies and cancer, *1849*(7), 861–870. https://doi.org/10.1016/j.bbagrm.2014.11.009.Stress

Apicco, D. J., Ash, P. E. A., Maziuk, B., LeBlang, C., Medalla, M., Abdullatif, A. Al, … Wolozin, B. (2018). Reducing the RNA binding protein TIA1 protects against tau-mediated neurodegeneration in vivo. *Nat Neurosci*, *21*(1), 72–80. https://doi.org/10.1038/nm.4346.Thermoneutral

Arimoto-Matsuzaki, K., Saito, H., & Takekawa, M. (2016). TIA1 oxidation inhibits stress granule assembly and sensitizes cells to stress-induced apoptosis. *Nature Communications*, *7*, 1–10. https://doi.org/10.1038/ncomms10252

Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H., & Takekawa, M. (2008). Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nature Cell Biology*, *10*(11), 1324–1332. https://doi.org/10.1038/ncb1791

Article, R., & Hwang, O. (2013). Role of Oxidative Stress in Parkinson's Disease. *Experimental Neurobiology*, *22*(1), 11–17.

Aulas, A., Fay, M. M., Lyons, S. M., Achorn, C. A., Kedersha, N., Anderson, P., & Ivanov, P. (2017). Stress-specific differences in assembly and composition of stress

granules and related foci. *J. Cell Sci.*, *130*(5), 927–937. https://doi.org/10.1016/j.ijfoodmicro.2006.06.014

Aulas, A., & Vande Velde, C. (2015). Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? *Frontiers in Cellular Neuroscience*, *9*(October), 1–13. https://doi.org/10.3389/fncel.2015.00423

Baguet, A., Degot, S., Cougot, N., Bertrand, E., Chenard, M.-P., Wendling, C., … Tomasetto, C. (2007). The exon-junction-complex-component metastatic lymph node 51 functions in stress-granule assembly. *Journal of Cell Science*, *120*(16), 2774–2784. https://doi.org/10.1242/jcs.009225

Balchin, D., Hayer-Hartl, M., & Hartl, F. U. (2016). In vivo aspects of protein folding and quality control. *Science*, *353*(6294). https://doi.org/10.1126/science.aac4354

Bandyopadhyay, U., Kaushik, S., Varticovski, L., & Cuervo, A. M. (2008). The Chaperone-Mediated Autophagy Receptor Organizes in Dynamic Protein Complexes at the Lysosomal Membrane. *Molecular and Cellular Biology*, *28*(18), 5747–5763. https://doi.org/10.1128/mcb.02070-07

Barber, G. N. (2005). The dsRNA-dependent protein kinase, PKR and cell death. *Cell Death and Differentiation*, *12*(6), 563–570. https://doi.org/10.1038/sj.cdd.4401643

Barreau, C., Paillard, L., & Osborne, H. B. (2005). AU-rich elements and associated factors: Are there unifying principles? *Nucleic Acids Research*, *33*(22), 7138–7150. https://doi.org/10.1093/nar/gki1012

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., … Zhang, J. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099

Bentmann, E., Haass, C., & Dormann, D. (2013). Stress granules in neurodegeneration - Lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. *FEBS Journal*, *280*(18), 4348–4370. https://doi.org/10.1111/febs.12287

Bley, N., Lederer, M., Pfalz, B., Reinke, C., Fuchs, T., Glaß, M., … Hüttelmaier, S. (2015). Stress granules are dispensable for mRNA stabilization during cellular stress. *Nucleic Acids Research*, *43*(4). https://doi.org/10.1093/nar/gku1275

Bounedjah, O., Desforges, B., Wu, T. Di, Pioche-Durieu, C., Marco, S., Hamon, L., … Pastré, D. (2014). Free mRNA in excess upon polysome dissociation is a scaffold for protein multimerization to form stress granules. *Nucleic Acids Research*, *42*(13), 8678–8691. https://doi.org/10.1093/nar/gku582

Brocker, C., Thompson, D. C., & Vasiliou, V. (2012). The role of hyperosmotic stress in inflammation and disease. *Biomol Concepts.*, *3*(4), 345–364. https://doi.org/10.1515/bmc-2012-0001.The

Bryant, C. D., & Yazdani, N. (2016). RNA-binding proteins, neural development and the addictions. *Genes, Brain and Behavior*, *15*(1), 169–186. https://doi.org/10.1111/gbb.12273

Buchan, J. R., Kolaitis, R., Taylor, J. P., & Parker, R. (2013). Eukaryotic Stress Granules Are Cleared by Autophagy and Cdc48 / VCP Function. *Cell*, *153*(7), 1461–1474. https://doi.org/10.1016/j.cell.2013.05.037

Buchan, J. R., Muhlrad, D., & Parker, R. (2008). P bodies promote stress granule assembly in Saccharomyces cerevisiae. *Journal of Cell Biology*, *183*(3), 441–455. https://doi.org/10.1083/jcb.200807043

Buchan, J. R., & Parker, R. (2009). Eukaryotic Stress Granules : The Ins and Out of Translation What are Stress Granules ? *Molecular Cell*, *36*(6), 932–948.

Buchan, J. R., Yoon, J.-H., & Parker, R. (2011). Stress-specific composition, assembly and kinetics of stress granules in Saccharomyces cerevisiae. *Journal of Cell Science*, *124*(2), 228–239. https://doi.org/10.1242/jcs.078444

Capes-davis, A., Theodosopoulos, G., Atkin, I., Drexler, H. G., Kohara, A., Macleod, R. A. F., … Freshney, R. I. (2010). Check your cultures ! A list of cross-contaminated or misidentified cell lines. *International Journal of Cancer*, *127*(1), 1–8. https://doi.org/10.1002/ijc.25242

Carrick, D. M., & Blackshear, P. J. (2007). Comparative expression of tristetraprolin (TTP) family member transcripts in normal human tissues and cancer cell lines. *Archives of Biochemistry and Biophysics*, *462*(2), 278–285. https://doi.org/10.1016/j.abb.2007.04.011

Carter, A. P. (2013). Crystal clear insights into how the dynein motor moves. *Journal of Cell Science*, *126*(3), 705–713. https://doi.org/10.1242/jcs.120725

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., … Hentze, M. W. (2012). Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, *149*(6), 1393–1406. https://doi.org/10.1016/j.cell.2012.04.031

Cencic, R., Carrier, M., Galicia-Vázquez, G., Bordeleau, M. E., Sukarieh, R., Bourdeau, A., … Pelletier, J. (2009). Antitumor activity and mechanism of action of the cyclopenta[b]benzofuran, silvestrol. *PLoS ONE*, *4*(4). https://doi.org/10.1371/journal.pone.0005223

Chen-plotkin, A. S., Lee, V. M., & Trojanowski, J. Q. (2010). TAR DNA-binding protein 43 in neurodegenerative disease. *Nat Rev Neurol.*, *6*(4), 211–220. https://doi.org/10.1038/nrneurol.2010.18.TAR

Chen, C. A., & Shyu, A. (1995). AU-rich elements: caracterization and importance in mRNA degradation. *Science*.

Cherkasov, V., Hofmann, S., Druffel-Augustin, S., Mogk, A., Tyedmers, J., Stoecklin, G., & Bukau, B. (2013). Coordination of translational control and protein homeostasis during severe heat stress. *Current Biology*, *23*(24), 2452–2462. https://doi.org/10.1016/j.cub.2013.09.058

Clarke, H. J., Chambers, J. E., Liniker, E., & Marciniak, S. J. (2014). Endoplasmic Reticulum Stress in Malignancy. *Cancer Cell*, *25*(5), 563–573. https://doi.org/10.1016/j.ccr.2014.03.015

Coates, P. J., Lorimore, S. A., & Wright, E. G. (2005). Cell and tissue responses to genotoxic stress. *The Journal of Pathology*, *205*(2), 221–235. https://doi.org/10.1002/path.1701

Collingridge, G., Olsen, R., Peters, J., & Spedding, M. (2009). A nomenclature for ligand-gated ion channels. *Neuropharmacology*, *56*(1), 2–5. https://doi.org/10.1016/j.neuropharm.2008.06.063.A

Cook, C. E., Bergman, M. T., Finn, R. D., Cochrane, G., Birney, E., & Apweiler, R. (2016). The European Bioinformatics Institute in 2016: Data growth and integration. *Nucleic Acids Research*, *44*(D1), D20–D26. https://doi.org/10.1093/nar/gkv1352

Cook, K. B., Kazan, H., Zuberi, K., Morris, Q., & Hughes, T. R. (2011). RBPDB: A database of RNA-binding specificities. *Nucleic Acids Research*, *39*(SUPPL. 1), 301–308. https://doi.org/10.1093/nar/gkq1069

Cooper-Knock, J., Hewitt, C., Highley, J. R., Brockington, A., Milano, A., Man, S., … Shaw, P. J. (2012). Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain*, *135*(3), 751–764. https://doi.org/10.1093/brain/awr365

Cooper, T. A., Wan, L., & Dreyfuss, G. (2009). RNA and Disease. *Cell*, *136*(4), 777–793.

Creacy, S. D., Beerbower, P. E., Smaldino, P. J., Vaughn, J. P., Routh, E. D., & Akman, S. A. (2017). A G-quadruplex DNA-affinity Approach for Purification of

Enzymatically Active G4 Resolvase1. *Journal of Visualized Experiments*, (121), 1–9. https://doi.org/10.3791/55496

Cruz, J. de la, Karbstein, K., & Jr, J. L. W. (2015). Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo. *Annu Rev Biochem*, *84*, 93–129. https://doi.org/10.1146/annurev-biochem-060614-033917.Functions

Curbelo, R. J., Loza, E., De Yébenes, M. J. G., & Carmona, L. (2014). Databases and registers: Useful tools for research, no studies. *Rheumatology International*, *34*(4), 447–452. https://doi.org/10.1007/s00296-014-2954-x

Dang, Y., Kedersha, N., Low, W. K., Romo, D., Gorospe, M., Kaufman, R., … Liu, J. O. (2006). Eukaryotic initiation factor 2α-independent pathway of stress granule induction by the natural product pateamine A. *Journal of Biological Chemistry*, *281*(43), 32870–32878. https://doi.org/10.1074/jbc.M606149200

Date, C. J. (2004). *Database Systems*.

David, C. J., & Manley, J. L. (2010). Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes & Development*, *24*, 2343–2364. https://doi.org/10.1101/gad.1973010.in

Dean, O. M., Buuse, M. Van Den, Berk, M., Copolov, D. L., Mavros, C., & Bush, A. I. (2011). N-acetyl cysteine restores brain glutathione loss in combined 2-cyclohexene-1-one and d -amphetamine-treated rats : Relevance to schizophrenia and bipolar disorder. *Neuroscience Letters*, *499*(3), 149–153. https://doi.org/10.1016/j.neulet.2011.05.027

DeGracia, D. J., Rudolph, J., Roberts, G. G., Rafols, J. A., & Wang, J. (2007). Convergence of Stress Granules and Protein Aggregates in Hippocampal CA1 At Later Reperfusion Following Global Brain Ischemia. *Neuroscience*, *146*(2), 562–572. https://doi.org/10.1002/cyto.a.20594.Use

Deo, R. C., Bonanno, J. B., Sonenberg, N., & Burley, S. K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell*, *98*(6), 835–845. https://doi.org/10.1016/S0092-8674(00)81517-2

Dewey, C. M., Cenik, B., Sephton, C. F., Johnson, B. A., Herz, J., & Yu, G. (2012). TDP-43 aggregation in neurodegeneration: are stress granules the key? *Brain Research*, *1462*(214), 16–25. https://doi.org/10.1016/j.brainres.2012.02.032.TDP-43

Dougherty, J. D., Tsai, W. C., & Lloyd, R. E. (2015). Multiple poliovirus proteins repress cytoplasmic RNA granules. *Viruses*, *7*(12), 6127–6140. https://doi.org/10.3390/v7122922

Edery, I., Humbelin, M., Darveau, A., Lee, K. a W., Milburn, S., Hershey, J. W. B., … Sonenberg, N. (1983). Involvement of eIF-4A in the cap recognition process. *Journal of Biological Chemistry*, *258*, 11398–11403.

Elden, A. C., Kim, H., Hart, M. P., Chen-plotkin, A. S., Johnson, S., Fang, X., … Gitler, A. D. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature*, *466*(7310), 1069–1075. https://doi.org/10.1038/nature09320.Ataxin-2

Emde, A., & Hornstein, E. (2014). miRNAs at the interface of cellular stress and disease. *The EMBO Journal*, *33*(13), 1428–1437. https://doi.org/10.15252/embj.201488142

Farny, N. G., Kedersha, N. L., & Silver, P. A. (2009). Metazoan stress granule assembly is mediated by P-eIF2α-dependent and -independent mechanisms. *Rna*, *15*(10), 1814–1821. https://doi.org/10.1261/rna.1684009

Fernández-Suárez, X. M., Rigden, D. J., & Galperin, M. Y. (2014). The 2014 Nucleic Acids Research Database Issue and an updated NAR online Molecular Biology Database Collection. *Nucleic Acids Research*, *42*(D1), 1–6. https://doi.org/10.1093/nar/gkt1282

Fernandez, E. J., & Lolis, E. (2002). Structure , Function , and Inhibition of Chemokines. *Annual Review of Pharmacology and Toxicology*, *42*(1), 469–499. https://doi.org/10.1146/annurev.pharmtox.42.091901.115838

Ferro, D., Yao, S., & Zarnescu, D. C. (2018). Dynamic duo – FMRP and TDP-43: Regulating common targets, causing different diseases. *Brain Research*, *1693*(April), 37–42. https://doi.org/10.1016/j.brainres.2018.04.034

Fittschen, M., Lastres-Becker, I., Halbach, M. V., Damrath, E., Gispert, S., Azizov, M., … Auburger, G. (2015). Genetic ablation of ataxin-2 increases several global translation factors in their transcript abundance but decreases translation rate. *Neurogenetics*, *16*(3), 181–192. https://doi.org/10.1007/s10048-015-0441-5

Fournier, M. J., Gareau, C., & Mazroui, R. (2010). The chemotherapeutic agent bortezomib induces the formation of stress granules. *Cancer Cell International*, *10*, 1–10. https://doi.org/10.1186/1475-2867-10-12

Fowler, A. M., & Alarid, E. T. (2007). Amping up estrogen receptors in breast cancer. *Breast Cancer Research*, *9*(4), 8–10. https://doi.org/10.1186/bcr1748

Fredericks, A. M., Cygan, K. J., Brown, B. A., & Fairbrother, W. G. (2015). RNA-Binding Proteins: Splicing Factors and Disease. *Biomolecules*, *5*, 893–909. https://doi.org/10.3390//biom5020893

Fujimura, K., Sasaki, A. T., & Anderson, P. (2012). Selenite targets eIF4E-binding protein-1 to inhibit translation initiation and induce the assembly of non-canonical stress granules. *Nucleic Acids Research*, *40*(16), 8099–8110. https://doi.org/10.1093/nar/gks566

Furukawa, Y., Kaneko, K., Matsumoto, G., Kurosawa, M., & Nukina, N. (2009). Cross-Seeding Fibrillation of Q/N-Rich Proteins Offers New Pathomechanism of Polyglutamine Diseases. *Journal of Neuroscience*, *29*(16), 5153–5162. https://doi.org/10.1523/JNEUROSCI.0783-09.2009

Galicia-Vázquez, G., Cencic, R., Robert, F., Agenor, A. Q., & Pelletier, J. (2012). A cellular response linking eIF4AI activity to eIF4AII transcription. *Rna*, *18*(7), 1373–1384. https://doi.org/10.1261/rna.033209.112

Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins. *Nature Reviews Genetics*, *15*(12), 829–845. https://doi.org/10.1038/nrg3813

Ghisolfi, L., Dutt, S., McConkey, M. E., Ebert, B. L., & Anderson, P. (2012). Stress Granules contribute to α-globin homeostasis in differentiating erythroid cells. *Biochem Biophys Res Commun*, *420*(4), 768–774. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L. M., & Anderson, P. (2004). Stress Granule Assembly Is Mediated by Prion-like Aggregation of TIA-1. *Molecular Biology of the Cell*, *15*, 5383–5398. https://doi.org/10.1091/mbc.E04

Giudice, G., Sánchez-Cabo, F., Torroja, C., & Lara-Pezzi, E. (2016). ATtRACT-a database of RNA-binding proteins and associated motifs. *Database*, *2016*(2), 1–9. https://doi.org/10.1093/database/baw035

Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters*, *582*(14), 1977–1986. https://doi.org/10.1016/j.febslet.2008.03.004

Gorgoni, B. (2004). The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: A developmental perspective. *Briefings in Functional Genomics and Proteomics*, *3*(2), 125–141. https://doi.org/10.1093/bfgp/3.2.125

Gorrini, C., Harris, I. S., & Mak, T. W. (2013). Modulation of oxidative stress as an

anticancer strategy. *Nature Reviews Drug Discovery*, *12*(12), 931–947. https://doi.org/10.1038/nrd4002

Grifo, J. A., Tahara, S. M., Morgan, M. A., Shatkin, A. J., & Merrick, W. C. (1983). New initiation factor activity required for globin mRNA translation. *The Journal of Biological Chemistry*, *258*(9), 5804–5810. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6853548

Gu, J., Chen, F., Iqbal, K., Gong, C. X., Wang, X., & Liu, F. (2017). Transactive response DNA-binding protein 43 (TDP-43) regulates alternative splicing of tau exon 10: Implications for the pathogenesis of tauopathies. *Journal of Biological Chemistry*, *292*(25), 10600–10612. https://doi.org/10.1074/jbc.M117.783498

Gu, J., Wu, F., Xu, W., Shi, J., Hu, W., Jin, N., … Liu, F. (2017). TDP-43 suppresses tau expression via promoting its mRNA instability. *Nucleic Acids Research*, *45*(10), 6177–6193. https://doi.org/10.1093/nar/gkx175

Hafner Česen, M., Pegan, K., Špes, A., & Turk, B. (2012). Lysosomal pathways to cell death and their therapeutic applications. *Experimental Cell Research*, *318*(11), 1245–1251. https://doi.org/10.1016/j.yexcr.2012.03.005

Haider, L., Fischer, M. T., Frischer, J. M., Bauer, J., Ho, R., Esterbauer, H., … Witztum, J. L. (2011). Oxidative damage in multiple sclerosis lesions. *BRAIN*, *134*, 1914–1924. https://doi.org/10.1093/brain/awr128

Halliwell, B. (2007). Oxidative stress and cancer: have we moved forward? *Biochem. J.*, *401*(1), 1–11. https://doi.org/10.1042/BJ20061131

Hanks, S. K., & Hunter, T. (2018). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *The FASEB Journal*, *9*(8), 576–596. https://doi.org/10.1096/fasebj.9.8.7768349

Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., & Ron, D. (2000). Stress-Induced Gene Expression in Mammalian Cells. *Molecular Cell*, *6*, 1099–1108. https://doi.org/S1097-2765(00)00108-8 [pii]

Heine, M., Heck, J., Ciuraszkiewicz, A., & Bikbaev, A. (2019). Dynamic compartmentalization of calcium channel signalling in neurons. *Neuropharmacology*, (February), 0–1. https://doi.org/10.1016/J.NEUROPHARM.2019.02.038

Hellen, C. U. T., & Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes and Development*, *15*(13), 1593–1612. https://doi.org/10.1101/gad.891101

Hilgers, R. I. E., Teixeira, D., & Parker, R. O. Y. (2006). Translation-independent inhibition of mRNA deadenylation during stress in. *RNA*, *12*(10), 1835–1845. https://doi.org/10.1261/rna.241006.stress

Hinnebusch, A. G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends in Biochemical Sciences*, *42*(8), 589–611. https://doi.org/10.1016/j.tibs.2017.03.004

Hofmann, S., Cherkasova, V., Bankhead, P., Bukau, B., & Stoecklin, G. (2012). Translation suppression promotes stress granule formation and cell survival in response to cold shock. *Molecular Biology of the Cell*, *23*(19), 3786–3800. https://doi.org/10.1091/mbc.e12-04-0296

Hogan, D. J., Riordan, D. P., Gerber, A. P., Herschlag, D., & Brown, P. O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. *PLoS Biology*, *6*(10), 2297–2313. https://doi.org/10.1371/journal.pbio.0060255

Homma, K., Suzuki, K., & Sugawara, H. (2011). The autophagy database: An all-inclusive information resource on autophagy that provides nourishment for research.

Nucleic Acids Research, *39*(SUPPL. 1), 986–990. https://doi.org/10.1093/nar/gkq995

Huang, C., Zhou, H., Tong, J., Chen, H., Liu, Y. J., Wang, D., … Xia, X. G. (2011). FUS transgenic rats develop the phenotypes of amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *PLoS Genetics*, *7*(3). https://doi.org/10.1371/journal.pgen.1002011

Hyman, A. A., Lee, H. O., Patel, A., Carra, S., Kopach, A., Alberti, S., … Maharana, S. (2017). An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *The EMBO Journal*, *36*(12), 1669–1687. https://doi.org/10.15252/embj.201695957

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., … Meyerson, M. (2012). Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell*, *150*(6), 1107–1120. https://doi.org/10.1016/j.cell.2012.08.029.Mapping

Ishigaki, S., Masuda, A., Fujioka, Y., Iguchi, Y., Katsuno, M., Shibata, A., … Ohno, K. (2012). Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. *Scientific Reports*, *2*, 1–8. https://doi.org/10.1038/srep00529

James, S. J., Cutler, P., Melnyk, S., Jernigan, S., Janak, L., Gaylor, D. W., & Neubrander, J. A. (2004). Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. *The American Journal of Clinical Nutrition*, *80*(6), 1611–1617.

Jiang, H.-Y., & Wek, R. C. (2005). GCN2 phosphorylation of eIF2alpha activates NF-kappaB in response to UV irradiation. *The Biochemical Journal*, *385*, 371–380. https://doi.org/10.1042/BJ20041164

Jiang, H., Wek, S. A., McGrath, B. C., Scheuner, D., Kaufman, R. J., Cavener, D. R., & Wek, R. C. (2003). Phosphorylation of the alpha subunit of eukaryotic initiation factor 2 is required for activation of NF-kappaB in response to diverse cellular stresses. *Molecular and Cellular Biology*, *23*(16), 5651–5663. https://doi.org/10.1128/MCB.23.16.5651

Jin, S., & White, E. (2007). Role of autophagy in cancer: Management of metabolic stress. *Autophagy*, *3*(1), 28–31. https://doi.org/10.4161/auto.3269

Johnson, E. C. B., Dammer, E. B., Duong, D. M., Yin, L., Thambisetty, M., Troncoso, J. C., … Seyfried, N. T. (2018). Deep proteomic network analysis of Alzheimer ' s disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease. *Molecular Neurodegeneration*, *13*(1), 1–22.

Jolla, L. (2006). Intracellular Water Homeostasis and the Mammalian Cellular Osmotic Stress Response. *Journal of Cellular Physiology*, *206*(1), 9–15. https://doi.org/10.1002/jcp.20445

Joseph, N., Zhang-James, Y., Perl, A., & Faraone, S. V. (2015). Oxidative Stress and ADHD: A Meta-Analysis. *J Atten Disord.*, *19*(11), 915–924. https://doi.org/10.1177/1087054713510354.Oxidative

Kahvejian, A., Svitkin, Y. V, Sukarieh, R., Boutchou, M. M., & Sonenberg, N. (2005). Mammalian poly (A) -binding protein is a eukaryotic translation initiation factor , which acts via multiple mechanisms. *Genes & Development*, *19*(1), 104–113. https://doi.org/10.1101/gad.1262905.(40S)

Kapp, L. D., & Lorsch, J. R. (2004). The Molecular Mechanics of Eukaryotic Translation. *Annual Review of Biochemistry*, *73*(1), 657–704. https://doi.org/10.1146/annurev.biochem.73.030403.080419

Karin, M., & Ben-neriah, Y. (2000). Phosphorylation Meets Ubiquitination: The Control

of NF-kB Activity. *Annual Review of Immunology*, 621–663. https://doi.org/DOI: 10.1146/annurev.immunol.18.1.621

Katzenellenbogen, J. A., Maggi, A., Cavailles, V., Dahlman-Wright, K., Gustafsson, J.-A., Muramatsu, M., … Korach, K. S. (2006). International Union of Pharmacology. LXIV. Estrogen Receptors. *Pharmacological Reviews*, *58*(4), 773–781. https://doi.org/10.1124/pr.58.4.8

Kaufman, R. J. (2004). Regulation of mRNA translation by protein folding in the endoplasmic reticulum. *Trends in Biochemical Sciences*, *29*(3), 152–158. https://doi.org/10.1016/j.tibs.2004.01.004

Kaushik, S., & Cuervo, A. (2012). Chaperone-mediated autophagy: a unique way to enter the lysosome world. *Trends Cell Biol.*, *22*(8), 407–417. https://doi.org/10.1016/j.tcb.2012.05.006.Chaperone-mediated

Kechavarzi, B., & Janga, S. C. (2014). Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biology*, *15*(1), 1–16. https://doi.org/10.1186/gb-2014-15-1-r14

Kedersha, N., & Anderson, P. (2009). *Regulation of Translation by Stress Granules and Processing Bodies*. *Progress in Molecular Biology and Translational Science* (1st ed., Vol. 90). Elsevier Inc. https://doi.org/10.1016/S1877-1173(09)90004-7

Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I. J., Stahl, J., & Anderson, P. (2002). Evidence That Ternary Complex (eIF2-GTP-tRNAiMet)– Deficient Preinitiation Complexes Are Core Constituents of Mammalian Stress Granules. *Molecular Biology of the Cell*, *13*, 195–210. https://doi.org/10.1091/mbc.01

Kedersha, N., Cho, M. R., Li, W., Yacono, P. W., Chen, S., Gilks, N., … Anderson, P. (2000). Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *Journal of Cell Biology*, *151*(6), 1257–1268. https://doi.org/10.1083/jcb.151.6.1257

Kedersha, N., Gupta, M., Li, W., Miller, I., & Anderson, P. (1999). RNA-binding Proteins TIA-1 and TIAR Link the Phosphorylation of eIF-2a to the Assembly of Mammalian Stress Granules. *The Journal of Cell Biology*, *147*(7), 1431–1441. https://doi.org/10.1083/jcb.147.7.1431

Kedersha, N., Ivanov, P., & Anderson, P. (2013). Stress granules and cell signaling: more than just a passing phase? *Trends in Biochemical Sciences*, *38*(10). https://doi.org/10.1037/a0018493.Understanding

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fitzler, M. J., … Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *Journal of Cell Biology*, *169*(6), 871–884. https://doi.org/10.1083/jcb.200502088

Kiebler, M. A., & Bassell, G. J. (2006). Neuronal RNA Granules: Movers and Makers. *Neuron*, *51*(6), 685–690. https://doi.org/10.1016/j.neuron.2006.08.021

Kim, W. J., Back, S. H., Kim, V., Ryu, I., & Jang, S. K. (2005). Sequestration of TRAF2 into Stress Granules Interrupts Tumor Necrosis Factor Signaling under Stress Conditions Sequestration of TRAF2 into Stress Granules Interrupts Tumor Necrosis Factor Signaling under Stress Conditions. *Molecular and Cellular Biology*, *25*(6), 2450–2462. https://doi.org/10.1128/MCB.25.6.2450

Kimball, S. R., Horetsky, R. L., Ron, D., Jefferson, L. S., & Harding, H. P. (2003). Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. *American Journal of Physiology-Cell Physiology*, *284*(2), C273–C284. https://doi.org/10.1152/ajpcell.00314.2002

Klionsky, D. J. (2008). Autophagy revisited: A conversation with Christian de Duve. *Autophagy*, *4*(6), 740–743. https://doi.org/10.4161/auto.6398

Klionsky, D. J., Cueva, R., & Yaver, D. S. (1992). Aminopeptidase I of Saccharomyces cerevisiae Is Localized to the Vacuole Independent of the Secretory Pathway. *The Journal of Cell Biology*, *119*(2), 287–299. https://doi.org/10.7868/s0026365616060161

Kobayashi, S. (2015). Choose Delicately and Reuse Adequately: The Newly Revealed Process of Autophagy. *Biological & Pharmaceutical BulletinBiological and Pharmaceutical Bulletin*, *38*(8), 1098–1103. https://doi.org/10.1248/bpb.b15-00096

Koch, B., Thumm, M., Wolf, D. H., Schlumpberger, M., Egner, R., Straub, M., & Veenhuis, M. (1994). Isolation of autophagocytosis mutants of Saccharomyces cerevisiae. *FEBS Letters*, *349*(2), 275–280. https://doi.org/10.1016/0014-5793(94)00672-5

Kolarova, M., García-Sierra, F., Bartos, A., Ricny, J., & Ripova, D. (2012). Structure and pathology of tau protein in Alzheimer disease. *International Journal of Alzheimer's Disease*, *2012*. https://doi.org/10.1155/2012/731526

Korth, H. F., & Silberschatz, A. (2011). *Database System Concepts*. https://doi.org/10.1145/253671.253760

Kroenke, D. M., & Auer, D. J. (2012). *Database Processing*.

Ku, D. (2007). Osmotic stress sensing and signaling in animals. *FEBS Journal*, *274*, 5781. https://doi.org/10.1111/j.1742-4658.2007.06097.x

Kültz, D., & Burg, M. (1998). Evolution of osmotic stress signaling via MAP kinase cascades. *The Journal of Experimental Biology*, *201*, 3015–3021.

Kwiatkowski, T. J., Bosco D A, Hosler, B. A., Cortelli, P., Jong, P. J. De, Yoshinaga, Y., & Haines, J. L. (2009). Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science*, *323*(February), 1205–1209.

Kwon, S. H., Zhang, Y., & Matthias, P. (2007). The deacetylase HDAC6 is an essential component of stress granules and plays a critical role in the cellular response to stress. *Genes and Development*, *21*, 3381–3394. https://doi.org/10.1101/gad.461107.lates

Lagier-Tourenne, C., Polymenidou, M., Hutt, K. R., Vu, A. Q., Baughn, M., Huelga, S. C., … Yeo, G. W. (2012). Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci*, *15*(11), 1488–1497. https://doi.org/10.1038/nn.3230.Divergent

Lauvrak, S. U., Munthe, E., Kresse, S. H., Stratford, E. W., & Namløs, H. M. (2013). Functional characterisation of osteosarcoma cell lines and identification of mRNAs and miRNAs associated with aggressive cancer phenotypes. *British Journal of Cancer*, *109*(8), 2228–2236. https://doi.org/10.1038/bjc.2013.549

Le, Y., Zhou, Y., Iribarren, P., & Wang, J. M. (2004). Chemokines and Chemokine Receptors: Their Manifold Roles in Homeostasis and Disease. *Cellular & Molecular Immunology*, *1*(2), 95–104.

Lee, J., Giordano, S., & Zhang, J. (2011). Autophagy, mitochondria and oxidative stress: cross-talk and redox signalling. *Biochemical Journal*, *441*(2), 523–540. https://doi.org/10.1042/bj20111451

Lee, M.-H., & Schedl, T. (2006). RNA-Binding Proteins. *WormBook*. https://doi.org/10.1016/B978-0-12-374984-0.00439-3

Lenzi, J., Santis, R. De, Turris, V. De, Morlando, M., Laneve, P., Calvo, A., … Chio, A. (2015). ALS mutant FUS proteins are recruited into stress granules in induced pluripotent stem cell-derived motoneurons. *Disease Models & Mechanisms*, *8*(7), 755–766. https://doi.org/10.1242/dmm.020099

Leung, A. K. L., Todorova, T., Ando, Y., & Chang, P. (2012). Poly(ADP-ribose) regulates post-transcriptional gene regulation in the cytoplasm. *RNA Biology*, *9*(5),

542–548.
Levin, E. R. (2005). Integration of the Extranuclear and Nuclear Actions of Estrogen. *Mol. Endocrinol.*, *19*(8), 1951–1959. https://doi.org/10.1146/annurev-cellbio-092910-154240.Sensory
Levine, B., Mizushima, N., & Virgin, H. W. (2011). Autophagy in immunity and inflammation. *Nature*, *469*(7330), 323–335. https://doi.org/10.1038/nature09782.Autophagy
Lewitzky, M., Simister, P. C., & Feller, S. M. (2012). Beyond "furballs" and "dumpling soups" - Towards a molecular architecture of signaling complexes and networks. *FEBS Letters*, *586*(17), 2740–2750. https://doi.org/10.1016/j.febslet.2012.04.029
Li, J. J., & Xie, D. (2014). RACK1, a versatile hub in cancer. *Oncogene*, *34*(15), 1890–1898. https://doi.org/10.1038/onc.2014.127
Li, Q., & Verma, I. M. (2002). NF-κB regulation in the immune system. *Nature Reviews Immunology*, *2*(10), 725–734. https://doi.org/10.1038/nri910
Li, W., Simarro, M., Kedersha, N., & Anderson, P. (2004). FAST Is a Survival Protein That Senses Mitochondrial Stress and Modulates TIA-1-Regulated Changes in Protein Expression FAST Is a Survival Protein That Senses Mitochondrial Stress and Modulates TIA-1-Regulated Changes in Protein Expression, *24*(24), 10718–10732. https://doi.org/10.1128/MCB.24.24.10718
Lin, Y., Protter, D. S. W., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins Article Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell*, *60*(2), 208–219. https://doi.org/10.1016/j.molcel.2015.08.018
Liu-Yesucevitz, L., Bilgutay, A., Zhang, Y. J., Vanderwyde, T., Citro, A., Mehta, T., … Wolozin, B. (2010). Tar DNA binding protein-43 (TDP-43) associates with stress granules: Analysis of cultured cells and pathological brain tissue. *PLoS ONE*, *5*(10), e13250. https://doi.org/10.1371/journal.pone.0013250
Lloyd, R. E. (2012). How Do Viruses Interact with Stress-Associated RNA Granules? *PLoS Pathogens*, *8*(6), 1002741–1002741. https://doi.org/10.1371/journal.ppat.1002741
López-Lastra, M., Rivas, A., & Barría, M. I. (2005). Protein synthesis in eukaryotes: The growing biological relevance of cap-independent translation initiation. *Biological Research*, *38*(2–3), 121–146. https://doi.org/10.4067/S0716-97602005000200003
López De Silanes, I., Fan, J., Yang, X., Zonderman, A. B., Potapova, O., Pizer, E. S., & Gorospe, M. (2003). Role of the RNA-binding protein HuR in colon carcinogenesis. *Oncogene*, *22*(46), 7146–7154. https://doi.org/10.1038/sj.onc.1206862
Loschi, M., Leishman, C. C., Berardone, N., & Boccaccio, G. L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. *Journal of Cell Science*, *122*(21), 3973–3982. https://doi.org/10.1242/jcs.051383
Lu, L., Han, A., & Chen, J. (2001). Translation Initiation Control by Heme-Regulated Eukaryotic Initiation Factor 2 α Kinase in Erythroid Cells under Cytoplasmic Stresses. *Mol Cell Biol*, *21*(23), 7971–7980. https://doi.org/10.1128/MCB.21.23.7971
Luchelli, L., Thomas, M. G., & Boccaccio, G. L. (2015). Synaptic control of mRNA translation by reversible assembly of XRN1 bodies. *Journal of Cell Science*, *128*(8), 1542–1554. https://doi.org/10.1242/jcs.163295
Lukong, K. E., Chang, K. wei, Khandjian, E. W., & Richard, S. (2008). RNA-binding proteins in human genetic disease. *Trends in Genetics*, *24*(8), 416–425. https://doi.org/10.1016/j.tig.2008.05.004
Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins : modular design

for efficient function. *Nat Rev Mol*, *8*(6), 479–490. https://doi.org/10.1038/nrm2178.RNA-binding

Mahboubi, H., Seganathy, E., Kong, D., & Stochaj, U. (2013). Identification of Novel Stress Granule Components That Are Involved in Nuclear Transport. *PLoS ONE*, *8*(6), e68356. https://doi.org/10.1371/journal.pone.0068356

Mahboubi, H., & Stochaj, U. (2017). Cytoplasmic stress granules: Dynamic modulators of cell signaling and disease. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1863*(4), 884–895. https://doi.org/10.1016/j.bbadis.2016.12.022

Malys, N., & McCarthy, J. E. G. (2011). Translation initiation: Variations in the mechanism can be anticipated. *Cellular and Molecular Life Sciences*, *68*(6), 991–1003. https://doi.org/10.1007/s00018-010-0588-z

Marrone, L., Poser, I., Casci, I., Japtok, J., Reinhardt, P., Janosch, A., … Sterneckert, J. L. (2018). Isogenic FUS-eGFP iPSC Repor ter Lines Enable Quantification of FUS Stress Granule Pathology that Is Rescued by Drugs Inducing Autophagy. *Stem Cell Reports*, *10*(2), 375–389. https://doi.org/10.1016/j.stemcr.2017.12.018

Masuda, K., Marasa, B., Martindale, J. L., Halushka, M. K., & Gorospe, M. (2009). Tissue- and age-dependent expression of RNA-binding proteins that influence mRNA turnover and translation. *Aging*, *1*(8), 681–698. Retrieved from http://impactaging.com/papers/v1/n8/full/100073.html

Matos, C. A., de Almeida, L. P., & Nóbrega, C. (2017). Proteolytic cleavage of polyglutamine disease-causing proteins: revisiting the toxic fragment hypothesis. *Current Pharmaceutical Design*, *23*, 753–775. https://doi.org/10.2174/1381612822666161227

Maziuk, B., Ballance, H. I., & Wolozin, B. (2017). Dysregulation of RNA Binding Protein Aggregation in Neurodegenerative Disorders. *Frontiers in Molecular Neuroscience*, *10*(April), 1–9. https://doi.org/10.3389/fnmol.2017.00089

Mazroui, R. (2002). Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. *Human Molecular Genetics*, *11*(24), 3007–3017. https://doi.org/10.1093/hmg/11.24.3007

Mazroui, R., Marco, S. Di, Kaufman, R. J., & Gallouzi, I.-E. (2007). Inhibition of the Ubiquitin-Proteasome System Induces Stress Granule Formation. *Molecular Biology of the Cell*, *18*(July), 2601–2618. https://doi.org/10.1091/mbc.E06

Mazroui, R., Sukarieh, R., Bordeleau, M.-E., Kaufman, R. J., Northcote, P., Tanaka, J., … Pelletier, J. (2006). Inhibition of Ribosome Recruitment Induces Stress Granule Formation Independently of Eukaryotic Initiation Factor 2 Phosphorylation. *Molecular Biology of the Cell*, *17*(October), 4212–4219. https://doi.org/10.1091/mbc.E06

McEwen, E., Kedersha, N., Song, B., Scheuner, D., Gilks, N., Han, A., … Kaufman, R. J. (2005). Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. *Journal of Biological Chemistry*, *280*(17), 16925–16933. https://doi.org/10.1074/jbc.M412882200

McKenney, R. J., Huynh, W., Tanenbaum, M. E., Bhabha, G., & Vale, R. D. (2014). Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes. *Science*, *345*(6194), 337–341. https://doi.org/10.1016/j.pain.2013.06.005.Re-Thinking

Milisav, I. (2011). *Cellular Stress Responses*. *Advances in Regenerative Medicine* (Vol. 10).

Mizushima, N., & Komatsu, M. (2011). Autophagy: Renovation of Cells and Tissues. *Cell*, *147*(4), 728–741. https://doi.org/10.1016/j.cell.2011.10.026

Mizushima, N., Ohsumi, Y., & Yoshimori, T. (2002). Autophagosome formation in mammalian cells. *Cell Structure and Function*, *27*(6), 421–429. https://doi.org/10.1247/csf.27.421

Mizushima, N., Yoshimori, T., & Ohsumi, Y. (2011). The Role of Atg Proteins in Autophagosome Formation. *Annual Review of Cell and Developmental Biology*, *27*(1), 107–132. https://doi.org/10.1146/annurev-cellbio-092910-154005

Moeller, B. J., Cao, Y., Li, C. Y., & Dewhirst, M. W. (2004). Radiation activates HIF-1 to regulate vascular radiosensitivity in tumors: Role of reoxygenation, free radicals, and stress granules. *Cancer Cell*, *5*(5), 429–441. https://doi.org/10.1016/S1535-6108(04)00115-1

Mokas, S., Mills, J. R., Garreau, C., Fournier, M.-J., Robert, F., Arya, P., … Mazroui, R. (2009). Uncoupling Stress Granule Assembly and Translation Initiation Inhibition. *Molecular Biology of the Cell*, *20*, 2673–2683. https://doi.org/10.1091/mbc.E08

Monahan, Z., Shewmaker, F., & Pandey, U. B. (2016). Stress granules at the intersection of autophagy and ALS. *Brain Research*, *1649*(Pt B), 189–200. https://doi.org/10.1016/j.brainres.2016.05.022

Moosavi, M. A., Haghi, A., Rahmati, M., Taniguchi, H., Mocan, A., Echeverría, J., … Atanasov, A. G. (2018). Phytochemicals as potent modulators of autophagy for cancer therapy. *Cancer Letters*, *424*, 46–69. https://doi.org/10.1016/j.canlet.2018.02.030

Morlando, M., Dini Modigliani, S., Torrelli, G., Rosa, A., Di Carlo, V., Caffarelli, E., & Bozzoni, I. (2012). FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *EMBO Journal*, *31*(24), 4502–4510. https://doi.org/10.1038/emboj.2012.319

Nath, A., Vetrovec, G. W., Cowley, M. J., Newton, M., DiSciascio, G., Mukharji, J., & Lewis, S. a. (1988). Glutamate Receptor Ion CHannels: Structure, Regulation, and Function. *Pharmacological Reveiws*, *14*(1), 37–40. https://doi.org/10.1124/pr.109.002451.405

Nunes, C., Mestre, I., Marcelo, A., Koppenol, R., Matos, C. A., & Nóbrega, C. (2019). MSGP: the first database of the protein components of the mammalian stress granules. *Database*, *2019*, 1–7. https://doi.org/10.1093/database/baz031

Ohn, T., Kedersha, N., Hickman, T., Tisdale, S., & Anderson, P. (2008). A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat Cell Biol.*, *10*(10), 1224–1231. https://doi.org/10.1038/ncb1783.A

Pabo, C. O., & Sauer, R. T. (1984). Protein-DNA Recognition. *Annual Review of Biochemistry*, *53*(1), 293–321. https://doi.org/10.1146/annurev.bi.53.070184.001453

Paddock, S. W. (1999). *Confocal Microscopy - Methods and Protocols* (Vol. 122).

Paglin, S., Hollister, T., Delohery, T., Hackett, N., McMahill, M., Sphicas, E., … Yahalom, J. (2001). A novel response of cancer cells to radiation involves autophagy and formation of acidic vesicles. *Cancer Research*, *61*(2), 439–444.

Pahl, H. L. (1999). Activators and target genes of Rel/NF-kappaB transcription factors. *Oncogene*, *18*(49), 6853–6866. https://doi.org/10.1038/sj.onc.1203239

Panas, M. D., Ivanov, P., & Anderson, P. (2016). Mechanistic insights into mammalian stress granule dynamics. *Journal of Cell Biology*, *215*(3), 313–323. https://doi.org/10.1083/jcb.201609081

Panas, M. D., Schulte, T., Thaa, B., Sandalova, T., Kedersha, N., Achour, A., & McInerney, G. M. (2015). Viral and Cellular Proteins Containing FGDF Motifs Bind G3BP to Block Stress Granule Formation. *PLoS Pathogens*, *11*(2), 1–22.

https://doi.org/10.1371/journal.ppat.1004659
Patel, V. P., & Chu, C. T. (2011). Nuclear transport , oxidative stress , and neurodegeneration. *Int J Clin Exp Pathol*, *4*(3), 215–229.
Paul, S., Dansithong, W., Figueroa, K. P., Scoles, D. R., & Pulst, S. M. (2018). Staufen1 links RNA stress granules and autophagy in a model of neurodegeneration. *Nature Communications*, *9*(3648). https://doi.org/10.1038/s41467-018-06041-3
Pelletier, J., Graff, J., Ruggero, D., & Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: A critical nexus for cancer development. *Cancer Research*, *75*(2), 250–263. https://doi.org/10.1158/0008-5472.CAN-14-2789.TARGETING
Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I., & Hellen, C. U. T. (2001). Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences*, *98*(13), 7029–7036. https://doi.org/10.1073/pnas.111145798
Petroni, G., Jensen, G. J., Ladinsky, M. S., McDowall, A. W., & Pilhofer, M. (2011). Microtubules in Bacteria: Ancient Tubulins Build a Five-Protofilament Homolog of the Eukaryotic Cytoskeleton. *PLoS Biology*, *9*(12), e1001213. https://doi.org/10.1371/journal.pbio.1001213
Phizicky, E. M., & Fields, S. (1995). Protein-Protein Interactions : Methods for Detection and Analysis. *Microbiological Reviews*, *59*(1), 94–123.
Pisareva, V. P., Pisarev, A. V., Komar, A. A., Hellen, C. U. T., & Pestova, T. V. (2008). Translation initiation on mammalian mRNAs with structured 5′- UTRs requires DExH-box protein DHX29. *Cell*, *135*(7), 1237–1250. https://doi.org/10.1016/j.cell.2008.10.037.Translation
Poblete-Durán, N., Prades-Pérez, Y., Vera-Otarola, J., Soto-Rifo, R., & Valiente-Echeverría, F. (2016). Who regulates whom? An overview of RNA granules and viral infections. *Viruses*, *8*(7), 1–28. https://doi.org/10.3390/v8070180
Poljšak, B., & Milisav, I. (2012). Clinical implications of cellular stress responses. *Bosnian Journal of Basic Medical Sciences*, *12*(2), 122–126. https://doi.org/10.17305/bjbms.2012.2510
Polymenidou, M., Lagier-tourenne, C., Hutt, K. R., Stephanie, C., Moran, J., Liang, T. Y., … Cleveland, D. W. (2011). Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci*, *14*(4), 459–468. https://doi.org/10.1038/nn.2779.Long
Porporato, P. E., Dhup, S., Dadhich, R. K., Copetti, T., & Sonveaux, P. (2011). Anticancer targets in the glycolytic metabolism of tumors: A comprehensive review. *Frontiers in Pharmacology*, *AUG*(August), 1–18. https://doi.org/10.3389/fphar.2011.00049
Pothof, J., Verkaik, N. S., Hoeijmakers, J. H. J., & Van Gent, D. C. (2009). MicroRNA responses and stress granule formation modulate the DNA damage response. *Cell Cycle*, *8*(21), 3462–3468. https://doi.org/10.4161/cc.8.21.9835
Protter, D. S. W., & Parker, R. (2016). Principles and Properties of Stress granules. *Trends Cell Biol.*, *26*(9), 668–679. https://doi.org/10.1016/j.tcb.2016.05.004.Principles
Rahbari, R., Sheahan, T., Modes, V., Collier, P., & Macfarlane, C. (2009). A novel L1 retrotransposon marker for HeLa cell line identification. *Biotechniques*, *46*(4), 277–284. https://doi.org/10.2144/000113089.A
Ralser, M., Albrecht, M., Nonhoff, U., Lengauer, T., Lehrach, H., & Krobitsch, S. (2005). An integrative approach to gain insights into the cellular function of human ataxin-2. *Journal of Molecular Biology*, *346*(1), 203–214. https://doi.org/10.1016/j.jmb.2004.11.024
Ralser, M., Nonhoff, U., Albrecht, M., Lengauer, T., Wanker, E. E., Lehrach, H., &

Krobitsch, S. (2005). Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways. *Human Molecular Genetics*, *14*(19), 2893–2909. https://doi.org/10.1093/hmg/ddi321

Ramond, A., Godin-ribuot, D., Ribuot, C., Totoson, P., Koritchneva, I., Cachot, S., … Joyeux-faure, M. (2011). Oxidative stress mediates cardiac infarction aggravation induced by intermittent hypoxia. *Fundamental & Clinical Pharmacology*, *27*(3), 252–261. https://doi.org/10.1111/j.1472-8206.2011.01015.x

Ravanidis, S., Kattan, F.-G., & Doxakis, E. (2018). Unraveling the Pathways to Neuronal Homeostasis and Disease: Mechanistic Insights into the Role of RNA-Binding Proteins and Associated Factors. *International Journal of Molecular Sciences*, *19*(8), 2280. https://doi.org/10.3390/ijms19082280

Richter, K., Haslbeck, M., & Buchner, J. (2010). The Heat Shock Response : Life on the Verge of Death. *Molecular Cell*, *40*(2), 253–266. https://doi.org/10.1016/j.molcel.2010.10.006

Rogelj, B., Easton, L. E., Bogu, G. K., Stanton, L. W., Rot, G., Curk, T., … Ule, J. (2012). Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Scientific Reports*, *2*, 1–10. https://doi.org/10.1038/srep00603

Ross Buchan, J. (2014). MRNP granules Assembly, function, and connections with disease. *RNA Biology*, *11*(8), 1019–1030. https://doi.org/10.4161/15476286.2014.972208

Ruggieri, A., Dazert, E., Metz, P., Hofmann, S., Bergeest, J., Mazur, J., … Bartenschlager, R. (2012). Dynamic Oscillation of Translation and Stress Granule Formation Mark the Cellular Response to Virus Infection. *Cell Host Microbe*, *12*(1), 1–24. https://doi.org/10.1016/j.chom.2012.05.013.Dynamic

Samson, A. L., Ho, B., Au, A. E., Schoenwaelder, M., Smyth, M. J., Bottomley, S. P., … Samson, A. L. (2016). Physicochemical properties that control protein aggregation also determine whether a protein is retained or released from necrotic cells. *Open Biology*, *6*(160098), 1–12.

Samson, A. L., Knaupp, A. S., Sashindranath, M., Borg, R. J., Au, A. E., Cops, E. J., … Medcalf, R. L. (2012). Nucleocytoplasmic Coagulation : An Injury-Induced Aggregation Event that Disulfide Crosslinks Proteins and Facilitates Their Removal by Plasmin. *Cell Reports*, *2*(4), 889–901. https://doi.org/10.1016/j.celrep.2012.08.026

Schueren, F., & Thoms, S. (2016). Functional Translational Readthrough: A Systems Biology Perspective. *PLoS Genetics*, *12*(8), 1–12. https://doi.org/10.1371/journal.pgen.1006196

Sebestyén, E., Singh, B., Miñana, B., Pagès, A., Mateo, F., Pujana, M. A., … Eyras, E. (2016). Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. *Genome Research*, *26*, 732–744. https://doi.org/10.1101/gr.242214.118

Segal, A. W. (2005). How Neutrophils Kill Microbes. *Annu Rev Immunol.*, *23*, 197–223. https://doi.org/10.1146/annurev.immunol.23.021704.115653.How

Seguin, S. J., Morelli, F. F., Vinet, J., Amore, D., De Biasi, S., Poletti, A., … Carra, S. (2014). Inhibition of autophagy, lysosome and VCP function impairs stress granule assembly. *Cell Death and Differentiation*, *21*(12), 1838–1851. https://doi.org/10.1038/cdd.2014.103

Sephton, C. F., Cenik, C., Kucukural, A., Dammer, E. B., Cenik, B., Han, Y., … Yu, G. (2011). Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *Journal of Biological Chemistry*, *286*(2), 1204–1215. https://doi.org/10.1074/jbc.M110.190884

Sharma, A., Lyashchenko, A. K., Lu, L., Nasrabady, S. E., Elmaleh, M., Mendelsohn, M., … Shneider, N. A. (2016). ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function. *Nature Communications*, *7*(10465), 1–14. https://doi.org/10.1038/ncomms10465

Shaw, G., & Kamen, R. (1986). A conserved adenine-uridine sequence from the 3' untranslated region of granulocyte-monocyte colony stimulating factor messenger RNA mediates selective messenger RNA degradation. *Cell*, *46*, 659–668. https://doi.org/10.1016/0092-8674(86)90341-7

Shen, S., Li, X.-F., Cullen, W. R., Weinfeld, M., & Chris Le, X. (2013). Arsenic Binding to Proteins. *Chemical Reviews*, *113*(10), 7769–7792. https://doi.org/10.1021/cr300015c

Shukla, S., & Parker, R. (2016). Hypo- and Hyper- Assembly Diseases of RNA-Protein Complexes. *Trends Mol Med*, *22*(7), 615–628. https://doi.org/10.1016/j.molmed.2016.05.005.Hypo-

Singh, N., Dhalla, A. K., Seneviratne, C., & Singal, P. K. (1995). Oxidative stress and heart failure. *Molecular and Cellular Biochemistry*, *147*(1–2), 77–81.

Slor, H. (1975). Differenttation between exonucleases and endonucleases and between haplotomic and diplotomic endonucleases using 3H-DNA-coated wells of plastic depression plates as substrate. *Nucleic Acids Research*, *2*(6), 897–903. https://doi.org/10.1093/nar/gkv1387

Somasekharan, S. P., El-Naggar, A., Leprivier, G., Cheng, H., Hajee, S., Grunewald, T. G. P., … Sorensen, P. H. (2015). YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. *Journal of Cell Biology*, *208*(7), 913–929. https://doi.org/10.1083/jcb.201411047

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, *136*(4), 731–745. https://doi.org/10.1016/j.cell.2009.01.042

Souquere, S., Mollet, S., Kress, M., Dautry, F., Pierron, G., & Weil, D. (2009). Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. *Journal of Cell Science*, *122*(20), 3619–3626. https://doi.org/10.1242/jcs.054437

Stefl, R., Skrisovska, L., & Allain, F. H. T. (2005). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Reports*, *6*(1), 33–38. https://doi.org/10.1038/sj.embor.7400325

Stoecklin, G., Stubbs, T., Kedersha, N., Wax, S., Rigby, W. F. C., Blackwell, T. K., & Anderson, P. (2004). MK2-induced tristetraprolin:14-3-3 Complexes prevent stress granule association and ARE-mRNA decay. *EMBO Journal*, *23*(6), 1313–1324. https://doi.org/10.1038/sj.emboj.7600163

Stöhr, N., Lederer, M., Reinke, C., Meyer, S., Hatzfeld, M., Singer, R. H., & Hüttelmaier, S. (2006). ZBP1 regulates mRNA stability during cellular stress. *Journal of Cell Biology*, *175*(4), 527–534. https://doi.org/10.1083/jcb.200608071

Suzuki, H., & Matsuoka, M. (2015). Overexpression of nuclear FUS induces neuronal cell death. *Neuroscience*, *287*, 113–124. https://doi.org/10.1016/j.neuroscience.2014.12.007

Takahashi, H., Iwata, T., Kitagawa, Y., Takahashi, R. H., Sato, Y., Wakabayashi, H., … Et Al. (1999). Increased levels of ε and isoforms of 14-3-3 proteins in cerebrospinal fluid in patients with Creutzfeldt-Jakob disease. *Clinical and Diagnostic Laboratory Immunology*, *6*(6), 983–985.

Takahashi, M., Higuchi, M., Makokha, G. N., Matsuki, H., Yoshita, M., Tanaka, Y., & Fujii, M. (2013). HTLV-1 Tax oncoprotein stimulates ROS production and

apoptosis in T cells by interacting with USP10. *Blood*, *122*(5), 715–725. https://doi.org/10.1182/blood-2013-03-493718

Tan, A. Y., Riley, T. R., Coady, T., Bussemaker, H. J., & Manley, J. L. (2012). TLS / FUS (translocated in liposarcoma / fused in sarcoma) regulates target gene transcription via single-stranded DNA response elements. *Pnas*, *109*(16), 6030–6035. https://doi.org/10.1073/pnas.1203028109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1203028109

Tavassoly, I. (2015). *Dynamics of Cell Fate Decision Mediated by the Interplay of Autophagy and Apoptosis in Cancer Cells*.

Tavassoly, I., Parmar, J., Shajahan-Haq, A. N., Clarke, R., Baumann, W. T., & Tyson, J. J. (2015). Dynamic modeling of the interaction between autophagy and apoptosis in mammalian cells. *CPT: Pharmacometrics and Systems Pharmacology*, *4*(4), 263–272. https://doi.org/10.1002/psp4.29

Taylor, J. P., Jr., R. H. B., & Cleveland, D. W. (2016). Decoding ALS: From Genes to Mechanism. *Nature*, *539*(7628), 197–206. https://doi.org/10.1038/nature20413.Decoding

Thedieck, K., Holzwarth, B., Prentzell, M. T., Boehlke, C., Kläsener, K., Ruf, S., … Baumeister, R. (2013). Inhibition of mTORC1 by Astrin and Stress Granules Prevents Apoptosis in Cancer Cells. *Cell*, *154*(4), 859–874. https://doi.org/10.1016/j.cell.2013.07.031

Thomas, A., Goumans, H., Amesz, H., Benne, R., & Voorma, H. O. (1979). A Comparison of the Initiation Factors of Eukaryotic Protein Synthesis from Ribosomes and from the Postribosomal Supernatant. *European Journal of Biochemistry*, *98*(2), 329–337. https://doi.org/10.1111/j.1432-1033.1979.tb13192.x

Thomas, M. G., Loschi, M., Desbats, M. A., & Boccaccio, G. L. (2011). RNA granules: the good, the bad and the ugly. *Cell Signal*, *23*(2), 324–334. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial

Tisdale, S., Dawra, N., Hickman, T., Emara, M. M., Hu, G.-F., Anderson, P., … Kedersha, N. (2010). Angiogenin-induced tRNA-derived Stress-induced RNAs Promote Stress-induced Stress Granule Assembly. *Journal of Biological Chemistry*, *285*(14), 10959–10968. https://doi.org/10.1074/jbc.m109.077560

Tollervey, J. R., Curk2, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., … Ule, J. (2011). Characterising the RNA targets and position-dependent splicing regulation by TDP-43; implications for neurodegenerative diseases. *Nat Neurosci*, *14*(4), 452–458. https://doi.org/10.1038/nn.2778

Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J. M., Bertrand, E., & Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. *Journal of Cell Biology*, *160*(6), 823–831. https://doi.org/10.1083/jcb.200212128

Treeck, B. Van, & Parker, R. (2019). Principles of Stress Granules Revealed by Imaging Approaches. *Cold Spring Harb Perspect Biol.*, *11*(2), 1–12. https://doi.org/10.1101/cshperspect.a033068

Tsuboi, S. (1992). Autophagy in Yeast Demonstrated with Proteinase-deficient Mutants and Conditions for its Induction. *The Journal of Cell Biology*, *119*(2), 201–311. https://doi.org/10.1055/s-2007-997145

Vale, R. D. (2003). The molecular motor toolbox for intracellular transport. *Cell*, *112*(4), 467–480. https://doi.org/10.1016/S0092-8674(03)00111-9

Valiente-Echeverría, F., Melnychuk, L., Vyboh, K., Ajamian, L., Gallouzi, I. E., Bernard, N., & Mouland, A. J. (2016). eEF2 and Ras-GAP SH3 domain-binding protein (G3BP1) modulate stress granule assembly during HIV-1 infection. *Nat Commun*,

5, 4819. https://doi.org/10.1038/ncomms5819.eEF2

Valko, M., Leibfritz, D., Moncol, J., Cronin, M. T. D., Mazur, M., & Telser, J. (2007). Free radicals and antioxidants in normal physiological functions and human disease. *International Journal of Biochemistry and Cell Biology*. https://doi.org/10.1016/j.biocel.2006.07.001

Vance, C., Rogelj, B., Hortobágyi, T., Vos, K. J. De, Lumi, A., Sreedharan, J., … Shaw, C. E. (2009). Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science*, *323*(5918), 1208–1211. https://doi.org/10.1126/science.1165942.Mutations

Vance, C., Scotter, E. L., Nishimura, A. L., Troakes, C., Mitchell, J. C., Kathe, C., … Shaw, C. E. (2013). ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. *Human Molecular Genetics*, *22*(13), 2676–2688. https://doi.org/10.1093/hmg/ddt117

Vanderweyde, T., Apicco, D. J., Youmans-kidder, K., Ash, P. E. A., Cook, C., Lummertz, E., … Wolozin, B. (2016). Interaction of tau with the RNA-Binding Protein TIA1 Regulates tau Pathophysiology and Toxicity. *Cell Rep*, *15*(7), 1455–1466. https://doi.org/10.1016/j.celrep.2016.04.045.Interaction

Vanderweyde, T., Youmans, K., Liu-Yesucevitz, L., & Wolozin, B. (2013). The Role Stress Granules and RNA Binding Proteins in Neurodegeneration. *Gerontology*, *59*(6), 997–1003. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Vilas-Boas, F. de A. S., da Silva, A. M., de Sousa, L. P., Lima, K. M., Vago, J. P., Bittencourt, L. F. F., … Barcelos, L. S. (2016). Impairment of stress granule assembly via inhibition of the eIF2alpha phosphorylation sensitizes glioma cells to chemotherapeutic agents. *Journal of Neuro-Oncology*, *127*(2), 253–260. https://doi.org/10.1007/s11060-015-2043-3

Waelter, S., Scherzinger, E., Wanker, E. E., Lehrach, H., Lurz, R., Boeddrich, A., & Lueder, G. (2001). Accumulation of Mutant Huntingtin Fragments in Aggresome-like Inclusion Bodies as a Result of Insufficient Protein Degradation. *Molecular Biology of the Cell*, *12*(5), 1393–1407. https://doi.org/10.1091/mbc.12.5.1393

Wang, W., Pan, L., Su, S. C., Quinn, E. J., Sasaki, M., Jimenez, J. C., … Tsai, L. (2013). Interaction of FUS and HDAC1 Regulates DNA Damage Response and Repair in Neurons. *Nat Neurosci*, *16*(10), 1383–1391. https://doi.org/10.1038/nn.3514.Interaction

Wang, W., Wang, L., Lu, J., Siedlak, S. L., Fujioka, H., Liang, J., … Wang, X. (2016). The Inhibition of TDP-43 Mitochondrial Localization Blocks Its Neuronal Toxicity. *Nat Med*, *22*(8), 869–878. https://doi.org/10.1038/nm.4130.The

Wang, Z. L., Li, B., Luo, Y. X., Lin, Q., Liu, S. R., Zhang, X. Q., … Qu, L. H. (2018). Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers. *Cell Reports*, *22*(1), 286–298. https://doi.org/10.1016/j.celrep.2017.12.035

Wek, R. C., Jiang, H.-Y., & Anthony, T. G. (2006). Coping with stress: eIF2 kinases and translational control. *Biochemical Society Transactions*, *34*(1), 7. https://doi.org/10.1042/BST20060007

Wek, S. A., Zhu, S., & Wek, R. C. (1995). The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase GCN2 interacts with tRNA and is required for activation in response to starvation for different amino acids. *Molecular and Cellular Biology*, *15*(8), 4497–4506. https://doi.org/10.1128/MCB.15.8.4497

Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R., & Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. *ELife*, *5*(Se), 1–25. https://doi.org/10.7554/eLife.18413

Wolozin, B. (2012). Regulated protein aggregation: Stress granules and

neurodegeneration. *Molecular Neurodegeneration*, *7*(56), 1–12. https://doi.org/10.1186/1750-1326-7-56

Wolozin, B. (2014). Physiological Protein Aggregation Run Amuck: Stress Granules and the Genesis of Neurodegenerative Disease. *Discov Med*, *17*(91), 47–52. https://doi.org/10.1016/j.chemosphere.2012.12.037.Reactivity

Wyttenbach, A., & Arrigo, A. P. (2009). The Role of Heat Shock Proteins during Neurodegeneration in Alzheimer's, Parkinson's and Huntington's Disease. *Heat Shock Proteins in Neural Cells*, 81–99. https://doi.org/10.1007/978-0-387-39954-6_7

Xie, Z., & Klionsky, D. J. (2007). Autophagosome formation: Core machinery and adaptations. *Nature Cell Biology*, *9*(10), 1102–1109. https://doi.org/10.1038/ncb1007-1102

Yamashita, M. M., Wesson, L., Eisenmant, G., & Eisenberg, D. (1990). Where metal ions bind in proteins. *Proc. Nadl. Acad. Sci. USA*, *87*(August), 5648–5652. Retrieved from http://www.pnas.org/content/pnas/87/15/5648.full.pdf

Yang, Z. J., Chee, C. E., Huang, S., & Sinicrope, F. A. (2011). The Role of Autophagy in Cancer: Therapeutic Implications. *Molecular Cancer Therapeutics*, *10*(9), 1533–1541. https://doi.org/10.1158/1535-7163.mct-11-0047

Yokoshi, M., Li, Q., Yamamoto, M., Okada, H., Suzuki, Y., & Kawahara, Y. (2014). Direct Binding of Ataxin-2 to Distinct Elements in 3' UTRs Promotes mRNA Stability and Protein Expression. *Molecular Cell*, *55*(2), 186–198. https://doi.org/10.1016/j.molcel.2014.05.022

Yu, C., York, B., Wang, S., Feng, Q., Xu, J., & O'Malley, B. W. (2007). An Essential Function of the SRC-3 Coactivator in Suppression of Cytokine mRNA Translation and Inflammatory Response. *Mol Cell*, *25*(5), 765–778. https://doi.org/10.1016/j.immuni.2010.12.017.Two-stage

Zhang, P., Fan, B., Yang, P., Temirov, J., Messing, J., Kim, H. J., & Taylor, J. P. (2019). Chronic optogenetic induction of stress granules is cytotoxic and reveals the evolution of ALS-FTD pathology. *ELife*, *8*, e39578.

Zinszner, H., Sok, J., Immanuel, D., Yin, Y., & Ron, D. (1997). TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *Journal of Cell Science*, *110*, 1741–1750. https://doi.org/10.1038/NATURE01031

Zlotnik, A., & Yoshie, O. (2012). The Chemokine Superfamily Revisited. *Immunity*, *36*(5), 705–716. https://doi.org/10.1016/j.immuni.2012.05.008.The

Zou, D., Ma, L., Yu, J., & Zhang, Z. (2015). Biological databases for human research. *Genomics, Proteomics and Bioinformatics*, *13*(1), 55–63. https://doi.org/10.1016/j.gpb.2015.01.006

ANNEXES

Annex A. Nunes,C., Mestre,I., Marcelo,A. *et al.* MSGP: the first database of the protein components of the mammalian stress granules. *Database* (2019) Vol. 2019

Database, 2019, 1–7
doi: 10.1093/database/baz031
Original article

Original article

MSGP: the first database of the protein components of the mammalian stress granules

Catarina Nunes[1,2], Isa Mestre[2], Adriana Marcelo[1,2,3], Rebekah Koppenol[1], Carlos A. Matos[1,2,3] and Clévio Nóbrega[1,2,3,4,*]

[1]Department of Biomedical Sciences and Medicine, University of Algarve, Faro, 8005-139, Portugal, [2]Centre for Biomedical Research, University of Algarve, Faro, 8005-139, Portugal, [3]Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, 3004-517, Portugal and [4]Algarve Biomedical Center, University of Algarve, Faro, 8005-139, Portugal

*Corresponding author: Tel.: +35191797250; Email: cdnobrega@ualg.pt

Citation details: Nunes,C., Mestre,I., Marcelo,A. *et al.* MSGP: the first database of the protein components of the mammalian stress granules. *Database* (2019) Vol. 2019: article ID baz031; doi:10.1093/database/baz031 Received 17 November 2018; Revised 14 January 2019; Accepted 11 February 2019

Abstract

In response to different stress stimuli, cells transiently form stress granules (SGs) in order to protect themselves and re-establish homeostasis. Besides these important cellular functions, SGs are now being implicated in different human diseases, such as neurodegenerative disorders and cancer. SGs are ribonucleoprotein granules, constituted by a variety of different types of proteins, RNAs, factors involved in translation and signaling molecules, being capable of regulating mRNA translation to facilitate stress response. However, until now a complete list of the SG components has not been available. Therefore, we aimer at identifying and linting in an open access database all the proteins described so far as components of SGs. The identification was made through an exhaustive search of studies listed in PubMed and double checked. Moreover, for each identified protein several details were also gathered from public databases, such as the molecular function, the cell types in which they were detected, the type of stress stimuli used to induce SG formation and the reference of the study describing the recruitment of the component to SGs. Expression levels in the context of different neurodegenerative diseases were also obtained and are also described in the database. The Mammalian Stress Granules Proteome is

available at https://msgp.pt/, being a new and unique open access online database, the first to list all the protein components of the SGs identified so far. The database constitutes an important and valuable tool for researchers in this research area of growing interest.

Introduction

Cells are exposed to different stress stimuli that they need to overcome ensuring cell survival. To manage stress, cells have several mechanisms ranging from repair pathways to apoptosis triggering, if cells fail to overcome the stress. Growing evidence suggests that a persistent cellular stress state might underlie an enhanced susceptibility to aging or aging-related diseases, like neurodegenerative disorders or cancer (1).

The assembly of stress granules (SGs) represents a conservative component of the cellular response to stress. SGs are ribonucleoprotein granules that appear when eukaryotic cells are exposed to certain types of stimuli such as endoplasmic reticulum stress,heat shock,hypoxia,arsenite, viral infection or overexpression of specific RNA-binding proteins (RBPs) (2). SGs are transiently formed upon cellular stress, and their disassembly occurs when the cellular stressor is removed. The canonical SG assembly pathway is triggered by the phosphorylation of eIF2α leading to the inhibition of translation, and thereby creating a pool of mRNAs stalled in translation initiation, translation initiation factors, RBPs and ribosomal units (3). SG assembly is key to cell survival as these foci inhibit apoptosis through reduction of reactive oxygen species (ROS), sequestration of signaling molecules and stabilization of mRNAs of antiapoptotic factors (4). Under stress conditions, global translation is reduced, and SGs are thought to function in the triage of repressed mRNAs, allowing a focused translation of proteins critical to overcome stress (5). Additionally, stalled mRNAs in SGs are protected from degradation during stress and can rapidly re-enter the translational pool once stress is overcome and they are released (6). Despite these important functions in translation and several others described, the complete functions of SGs are not yet understood.

The molecular composition of SG core is based in stalled mRNA transcripts, poly(A) mRNAs, RBPs, translation initiation factors, proteins with predicted lowcomplexity domains and small (40S) ribosomal units (7). Due to their frequent presence in SGs, some proteins, are commonly used as SG markers,including several eukaryotic initiation factors, poly(A)-binding protein 1 (PABP1), Tcell intracellular antigen 1 (TIA-1), TIA-1-related protein (TIAR), Ras GTPase-activating protein-binding protein 1 (G3BP1) and ataxin-2 (8). Nevertheless, SG composition changes during the stress response and is also different according to the type of stress or cell (9). In fact, recently, it was found that ∼20% of SG components diversity is dependent on the stress and the cell type (10).

Growing and recent evidence implicates SGs in the context of human disease, namely in cancer (2) and in neurodegenerative disorders (11). For example, in cancer, SGs were found in different tumors with different histological origins (12–14). In the same line, in Alzheimer's disease (AD) several SG components accumulate in affected cells and colocalize with pathogenic tau (15, 16). We also showed that, in the context of another neurodegenerative disease - Machado–Joseph, the SG component ataxin-2 is downregulated, contributing decisively to the pathology, whereas its overexpression ameliorates the disease phenotype (17). On the other hand, antisense oligonucleotidesmediated ATXN2 silencing was successful in reducing neuropathological abnormalities in spinocerebellar ataxia type 2 and amyotrophic lateral sclerosis animal models (18, 19). Additionally, SGs could also be implicated in the normal aging process, as a reduction in the expression of several SG components with age, especially RBPs, has been described (20).

A database consists in a storage of information that can be easily accessed and that is regularly managed and updated. Therefore, databases serve as an important tool for research and, accordingly, the number of databases has been increasing in the past years (21). Despite the growing interest in SG research and several reviews on the topic, there is still a lack of resources for their study. There are several important databases on RBPs, focusing on different aspects of their structure or function, although they do not address the RBPs' role/presence in SGs (22, 23). SGs were originally described in tomato cell lines submitted to heat shock (24), and since then several studies demonstrated the recruitment of different proteins to SGs. However, the complete list of SG components is unknown.Thus,we generated electronic resources in the form of Excel-based databases/ tables containing all the protein components recruited to SGs that have been described so far. These data were the basis for the development of an online database available at https://msgp.pt/, which we now present. The database curates general information about all the protein components of SGs described so far in mammalian cells. The platform provides a new and unique resource for the SG research field, collecting and storing for the first time and in the same place all the information on the SG protein components.

Material and methods

Components identification and curation

We curated the published literature available in different databases (like for example PubMed) covering all the SG protein components described in studies using mammalian cells. Several keyword combinations were used, such as stress granules AND mammals or stress granules AND sodium arsenite. Each study describing the recruitment of different proteins to SGs was double checked, and the type of cell, stimulus used and effective recruitment of the component to the SGs were annotated and confirmed. Additionally, for all the identified and validated SG components several details were gathered from public databases, including the protein abbreviation, gene ID, chromosomal location, Uniprot ID, molecular function, subcellular localization, original study describing its recruitment to SGs (along with the cell type and stimulus used), RBP classification [according to (25)], identity as autophagy-related proteins [according to (26)] and the OMIM details for their possible implication in human genetic disorders. All these details were gathered in the form of Excel-based databases/tables.

Gene expression data analysis

The GEO Expression Omnibus public database was used to find studies describing gene expression data in different neurodegenerative diseases. From the found studies, three were chosen based on the high number of sampled individuals, as well as on the type of neurodegenerative disease studied.

Expression profiles for all the identified and curated SG components were extracted from a transcriptome data set of human brain biopsy tissue sample, covering subjects with AD, Huntington's disease (HD) and healthy controls [GSE33000; (27)]; subjects with Parkinson's disease (PD) and healthy controls (GSE28894); and subjects with amyotrophic lateral sclerosis (ALS) and healthy controls [GSE4595; (28)]. We analyzed the original expression data using the Geo2R web tool (http://www.ncbi.nlm.nih.gov/ geo/geo2r/) comparing the different groups: AD versus controls, HD versus controls, PD versus controls and ALS versus controls. An adjusted $P < 0.05$ accessed the SG components whose expression was statistically different between groups.

Adjustments were made to correct the occurrence of false positive results, using the Benjamini and Hochberg false discovery rate method.

Software tools and database implementation

The Mammalian Stress Granules Proteome (MSGP) is an online and open access database. The website was implemented using the Wordpress system, and the database was build using 'custom fields', based on the open source tool 'Elementor Page Builder'.The database includes,also,other important features like custom listing profile for each protein, custom fields with editing capability for each protein, highly customized GeneID cards, protein listing quick view, breadcrumbs navigation, custom dashboard for front and end users and customized and multiple IDs for each protein. We also included >50 widgets ready to use on the database (keeping in mind its future expansion), integrated in a clean system, compatible with PHP version 5.5+, and we used minified and combined assets to reduce the amount of http requests and enhance load time and site performance. The system was built using Vue.js JavaScript framework and we programmed clean and well-structured code, to facilitate access to the data. The MSGP platform was also conceived to be responsive, working on all types of devices (mobile phones, tablets, computers etc.) and integrating the future functionality of user sign in/registration.

Results

Database structure

The MSGP database was developed in a highly customized way,allowingthe introduction of several important features and future uses. The structure and navigation were planned to work in a very intuitive and user-friendly manner. Moreover, a 'How to use' section where database users can find a tutorial showing the search process as well as the system navigation, especially in each SG component-specific page, was also created. The MSGP database contains primarily the 464 proteins identified as components of the SGs,which can be retrieved through different pages from the database website or from a general search. The structure of the database is described in Figure 1, providing three major sections: All proteins, grouping the information and details on the 464 proteins identified; the RBP section, where all SG components classified as RBPs are grouped; and the Autophagy section, where all SG components that belong to the autophagy pathway are grouped. Additionally, the database has a page 'Protein Index' where all the SG components are listed alphabetically. The database allows further exploring of the identified and curated components, in the 'Featured Proteins' section or in the 'New Proteins' section, where the proteins most recently included in the database are added. Furthermore, the database allows exploring the listed proteins according to different tags, such as sodium arsenite, U-2 OS or Hela cells.

Database content

For each one of the 464 identified proteins a set of details can also be found, including for example its molecular function,the complete protein name or its subcellular localization.Moreover,in each component we detail the original study describing its recruitment to SGs, as well as the type of stimulus used to induce SG assembly and the type of cell employed in that study. Due to the importance of RBPs to the nucleation and formation of SGs we also

detail if the identified component is an RBP or not. Interestingly and as expected, from the 464 proteins identified 252 (54%) are classified as RBPs according to Castello *et al.*'s (25)

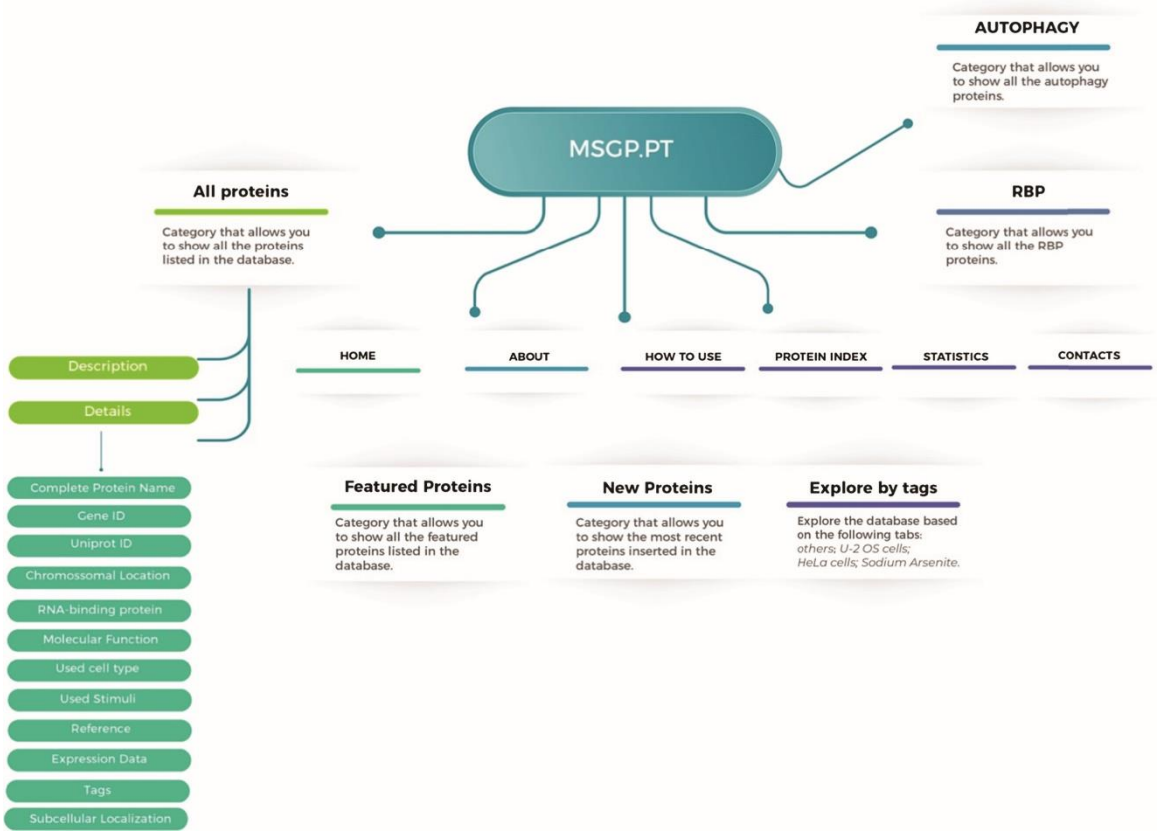


Figure 1. Structure of the MSGP database, depicting the main sections, different filters, tags and forms to explore and retrieve the information stored in the database.

study and to the Uniprot database. In line with this data, we performed a molecular function analysis of the 464 SG components (using the Protein Analysis Through Evolutionary Relationships available at http://www.pantherdb. org), which revealed that majority of the identified proteins have a binding or catalytic activity (Figure 2).

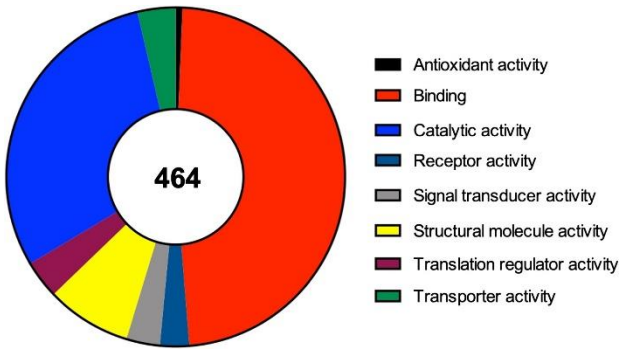


Figure 2. Molecular function of the 464 proteins currently identified as components of mammalian SGs, according to a gene list analysis by the PANTHER classification system.

Online interface

The online interface of the MSGP database homepage has a 'Search' form and three main tabs, 'All Proteins', 'RBP' and 'Autophagy' (Figure 3A). The 'Search' form allows a free text exploration of possible components of the database; the tab 'All proteins' lists all the proteins listed in the database; the tab 'RBP' lists the proteins classified in the database as RBPs; the tab 'Autophagy' groups all the proteins involved in this pathway (26) that are recruited to SGs. In the search form, four categories can be selected: (i) All Proteins, (ii) RBP, (iii) OMIM disease and (iv) Autophagy (Figure 3B). The first option allows a search in the entire database,whereasthe three otheroptions narrow the search to the SG components that are RBPs, linked to a human disease or to autophagy, respectively. The database landing page also has additional filters and tags that can be used to refine the listed proteins in the database (Figure 3C). For example, the 'U-2 OS cells' tab at the end of the page groups the SG components that were identified in these cells. As already mentioned, all the proteins in the database are listed in alphabetical order in the page 'Protein Index' to facilitate the search and to collect in the same page all the components of the SGs in the database (Figure 3A).

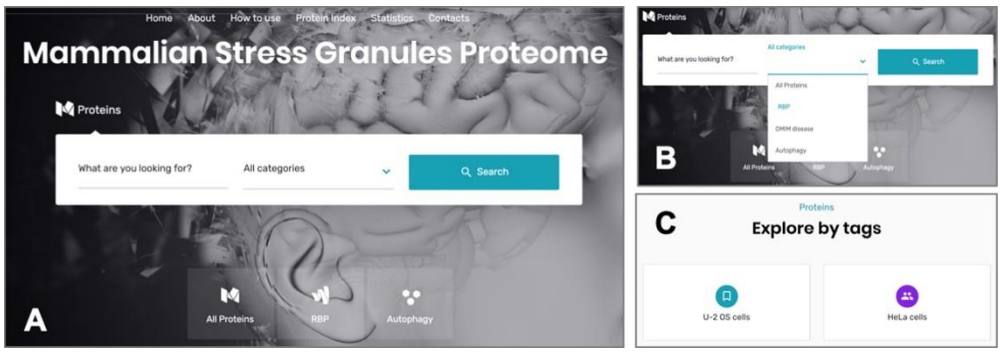


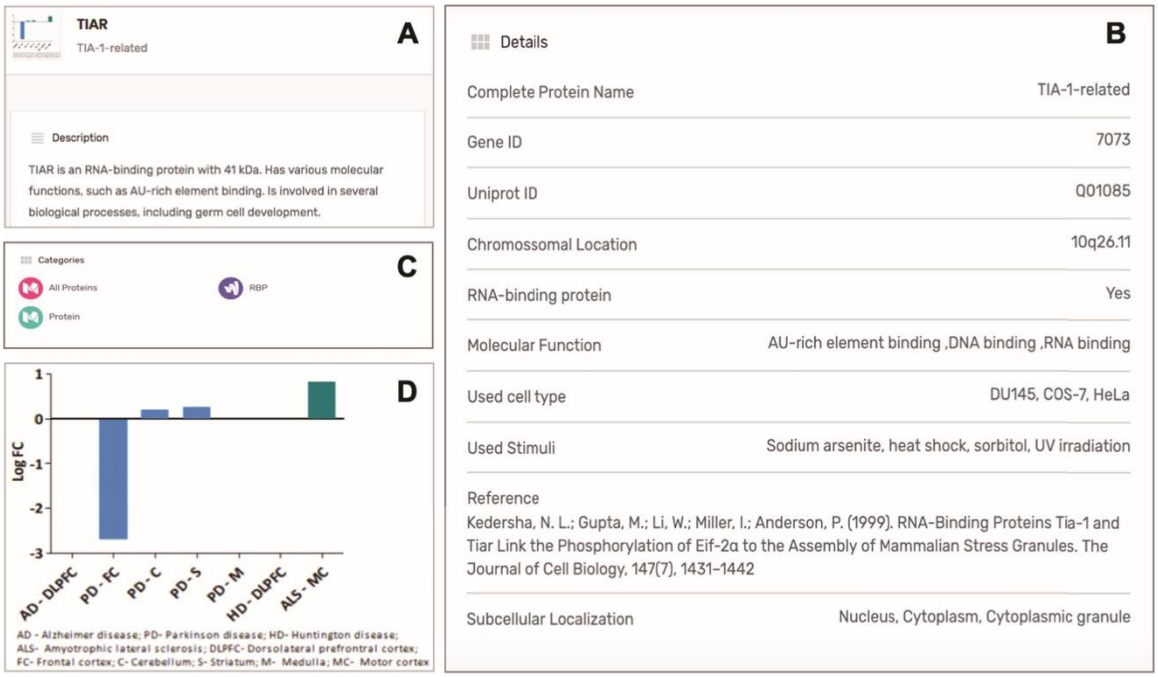
Figure 3. (A) The online interface of the database, detailing the landing page with the search form and the different pages. **(B)** The different options to narrow the search. **(C)** The possibility of accessing the listed components through different tags.

Searching and search results

The components listed in the database can be searched using the most common alias for the protein, which could be found in databases such as Uniprot or GeneCards. For example, for the core SG component GTPase-activating protein (SH3 domain) binding protein 1 the search should be performed using the alias 'G3BP1'. If no selection is made, the search is performed in the 'All Proteins' category (Figure 3B). Each component must be searched alone, as the use of 'AND', 'OR' and 'NOT' is not recognized by the database. If several SG components are listed with similar names, as in case as they belong for example to the same family, the search will list all those proteins. For example, the search for 'PABP' in the database will result in three hits: PABP1, PABP3 and PABP4.

Specific component pages

Each protein listed as a SG component in the MSGP database has a specific page where the different information details are listed. Independently from the form used to find a specific protein in the database (search, tabs, filters, index or general sections), the individual page for each protein is the same (Figure 4A). The complete name of the protein, the gene and Uniprot IDs, the chromosomal location, RBP classification (yes or no), the molecular function, OMIM details and subcellular localization are detailed for each component (Figure 4B). Each page also describes the category/categories where that specific component was classified (Figure 4C).



Importantly, the original study describing the recruitment of that protein to SGs is referred, as well as the type of stimulus used to induce SGs assembly and the type of cell used in the study. In each SG component page it is also possible to find the gene expression values for that component in the context of different neurodegenerative diseases (Figure 4D).

Figure 4. (A) Information details described in the specific page for the SG component TIAR, including a brief description of its molecular function. **(B)** Several details for the component are listed in the database, including the study describing the recruitment of this protein to SGs. **(C)** Each SG components is also included in specific categories, which are also displayed in its specific page. **(D)** For each SG component its expression levels in the context of neurodegenerative diseases are also described in the form of a graph, based on a differential expression analysis.

Expression data

The expression data for each of the SG components listed was extracted from different published studies and available in open access databases. The analysis compared the expression levels in patients with different neurodegenerative diseases and healthy controls. The differential expression level for each SG component in the different neurodegenerative diseases was plotted into a graphic and isalsoapartofthecomponent-specificpage(Figure 4D).As mentioned, several studies implicate SGs in different neurodegenerative diseases. Moreover, most of SG components are RBPs, which in the context of neurons are involved in different processes such as alternative splicing, transport, localization and stability and translation of RNAs (29). Thus, alterations in their expression may underlie or have impact on the

neurodegenerative pathogenesis. In fact, the analysis of the differential expression of the 464 SG components detected that 380 components have the expression altered in the brain of AD patients (Figure 5A). Similarly, in the brain of HD patients, 395 have their expression significantly altered. From these, 191 SG components have their expression commonly altered in AD and HD, with 90 having the expression increased and 101 repressed (Figure 5B).

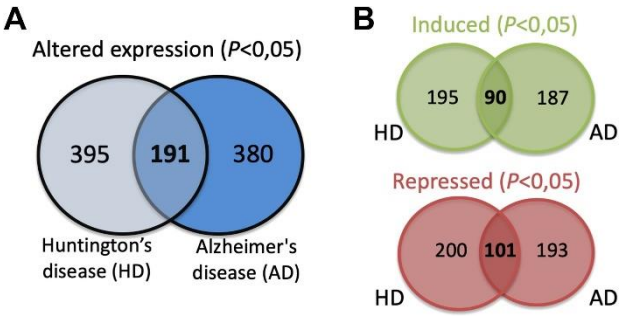


Figure 5. (A) SG components whose expression is significantly altered in the brain of AD and HD patients and commonly in both diseases. **(B)** Details on the number of SGs protein components whose expression is significantly induced or repressed in each disease and in both diseases.

Components identification and curation Users interaction and future updates

At the moment the database does not have a page for users to submit new SG components; however, researchers could send that information to the contacts in the database and shortly it will be updated. Nevertheless, there will be a continuous updating for new SG components from the published studies indexed in PubMed. The entire data set of the database is also available for any researcher upon request by email to cdnobrega@ualg.pt. The future plans for updating the database include adding information on the gene expression levels of SG components for different types of cancer. Importantly, we will continue to complete each protein details with relevant information, especially SG-related, as for example the number of studies where that particular component was described. We also consider to include other SG components besides proteins, such as miRNAs or mRNAs.

Conclusions

The MSGP database is the first tool cataloging all the SGs' protein components described so far. Moreover, it also collects several details about each component, thus providing an important tool for researchers in this area. The growing interest in the SG field and their implication in different human diseases make this database actual and opportune. Furthermore, the MSGP database has the possibility of being continuously updated as more components are described in SGs, and also of being expanded by adding more information and details, for example detailing the expression levels of these components in the context of different types of cancer. The database will constitute an important asset for SG research and will be continuously improved, based on our already defined plans and on the feedback from users.

Funding

This study was funded by the Portuguese Science and Technology Foundation (FCT) project (ALG-01- 0145-FEDER-29480) "SeGrPolyQ". The manuscript publication was also supported by the Algarve Biomedical Center (ABC) and the Município de Loulé. CN's lab is funded by the French Muscular Dystrophy Association and Ataxia UK. AM is supported by a Ph.D. fellowship from FCT (SFRH/BD/133192/2017).

Conflict of interest. None declared.

References

1. Poljsak,B. and Milisav,I. (2012) Clinical implications of cellularˇ stress responses. *Bosn. J. Basic Med. Sci.*, 12, 122–126.
2. Anderson,P., Kedersha,N. and Ivanov,P. (2015) Stress granules, P-bodies and cancer. *Biochim. Biophys. Acta*, 1849, 861–870.
3. Thomas,M.G.,Loschi,M.,Desbats,M.A.*etal.*(2011)RNAgranules: the good, the bad and the ugly. *Cell Signal.*, 23, 324–334.
4. Gareau,C., Fournier,M.J., Filion,C. *et al.* (2011) p21(WAF1/CIP1) upregulation through the stress granuleassociated protein CUGBP1 confers resistance to bortezomibmediated apoptosis. *PLoS One*, 6, e20254.
5. Anderson,P.and Kedersha,N.(2006) RNA granules.*J.Cell Biol.*, 172, 803–808.
6. Chantarachot,T. and Bailey-Serres,J. (2018) Polysomes, stress granules, and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function. *Plant Physiol.*, 176, 254–269.
7. Khong,A., Matheny,T., Jain,S. *et al.* (2017) The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol. Cell*, 68, 808–820 e805.
8. Carmo-Silva,S., Nobrega,C., Pereira de Almeida,L. *et al.* (2017) Unraveling the role of ataxin-2 in metabolism. *Trends Endocrinol. Metab.*, 28, 309–318.
9. Vanderweyde,T., Youmans,K., Liu-Yesucevitz,L. *et al.* (2013) Role of stress granules and RNA-binding proteins in neurodegeneration: a mini-review. *Gerontology*, 59, 524–533.
10. Markmiller,S., Soltanieh,S., Server,K.L. *et al.* (2018) Contextdependent and disease-specific diversity in protein interactions within stress granules. *Cell*, 172, 590–604.
11. Wolozin,B. (2012) Regulated protein aggregation: stress granules and neurodegeneration. *Mol. Neurodegener.*, 7, 56.
12. Vilas-Boas Fde,A., da Silva,A.M., de Sousa,L.P. *et al.* (2016) Impairment of stress granule assembly via inhibition of the eIF2alpha phosphorylation sensitizes glioma cells to chemotherapeutic agents. *J. Neurooncol.*, 127, 253–260.
13. Adjibade,P., St-Sauveur,V.G., Quevillon Huberdeau,M. *et al.* (2015) Sorafenib, a multikinase inhibitor, induces formation of stress granules in hepatocarcinoma cells. *Oncotarget*, 6, 43927–43943.
14. Somasekharan,S.P., El-Naggar,A., Leprivier,G. *et al.* (2015) YB1 regulates stress granule formation and tumor progression by translationally activating G3BP1. *J. Cell Biol.*, 208, 913–929.
15. Vanderweyde,T., Yu,H., Varnum,M. *et al.* (2012) Contrasting pathology of the stress granule proteins TIA-1 and G3BP in tauopathies. *J. Neurosci.*, 32, 8270–8283.
16. Ash,P.E., Vanderweyde,T.E., Youmans,K.L. *et al.* (2014) Pathological stress granules in Alzheimer's disease. *Brain Res.*, 1584, 52–58.
17. Nóbrega,C., Carmo-Silva,S., Albuquerque,D. *et al.* (2015) Reestablishing ataxin-2 downregulates translation of mutant ataxin-3 and alleviates Machado–Joseph disease. *Brain*, 138, 3537–3554.
18. Scoles,D.R., Meera,P., Schneider,M.D. *et al*. (2017) Antisense oligonucleotide therapy for spinocerebellar ataxia type 2. *Nature*, 544, 362–366.
19. Becker,L.A., Huang,B., Bieri,G. *et al.* (2017) Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP43 mice. *Nature*, 544, 367–371.
20. Masuda,K., Marasa,B., Martindale,J.L. *et al.* (2009) Tissueand age-dependent expression of RNA-binding proteins that influence mRNA turnover and translation. *Aging (Albany NY)*, 1, 681–698.
21. Curbelo,R.J., Loza,E., de Yébenes,M.J. *et al.* (2014) Databases and registers: useful tools for research, no studies. *Rheumatol. Int.*, 34, 447–452.
22. Giudice,G., Sanchez-Cabo,F., Torroja,C. *et al.* (2016) ATtRACT—a database of RNA-binding proteins and associated motifs. *Database (Oxford)*, 2016.
23. Cook,K.B., Kazan,H., Zuberi,K. *et al.* (2011) RBPDB: a database of RNA-binding specificities. *Nucleic Acids Res.*, 39, D301–D308.
24. Nover,L., Scharf,K.D. and Neumann,D. (1983) Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. *Mol. Cell. Biol.*, 3, 1648–1655.
25. Castello,A., Fischer,B., Eichelbaum,K. *et al*. (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*, 149, 1393–1406.

26. Homma,K., Suzuki,K. and Sugawara,H. (2011) The autophagy database: an all-inclusive information resource on autophagy that provides nourishment for research. *Nucleic Acids Res.*, 39, D986–D990.
27. Narayanan,M., Huynh,J.L., Wang,K. *et al.* (2014) Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases. *Mol. Syst. Biol.*, 10, 743.
28. Lederer,C.W., Torrisi,A., Pantelidou,M. *et al.* (2007) Pathways and genes differentially expressed in the motor cortex of patients with sporadic amyotrophic lateral sclerosis. *BMC Genomics*, 8, 26.
29. Bryant,C.D.and Yazdani,N.(2016) RNA-binding proteins,neural development and the addictions. *Genes Brain Behav.*, 15, 169–186.

Annex B. Stress granules' proteins. All proteins identified as components of Stress Granules.

ACTBL2	CHP1	DZIP1	GFPT1
ACTR1A	CIRP	EDC4	GLE1
ACTR1B	CIT	EIF2A	GNB2
ADAR1	CLIC4	EIF2AK1	GRB7
Ago1	CNBP	EIF2AK2	GSPT1
Ago2	CNN3	EIF2S2	H1F0
AKAP9	CNOT1	EIF3A	H1FX
ALDH18A1	CORO1B	EIF3B	H2AFV
ANG	CPEB1	EIF3D	HABP4
ANP32E	CPSF3	EIF3E	HDAC6
ANXA1	CPSF6	EIF3F	HELZ
ANXA6	CPSF7	EIF3G	HELZ2
ANXA7	CRYAB	EIF3H	HMGA1
APEX1	CSDE1	EIF3I	HMGB3
APOBEC3G	CSE1L	EIF3J	HMGN1
ARPC1B	CSTF1	EIF3K	HNRNPA2B1
ATAD2	CTNNA2	EIF3L	HNRNPA3
ATAD3A	CTNND1	EIF3M	HNRNPAB
ATP2C1	CTTNBP2NL	EIF4A1	HNRNPD
ATP5F1A	CWC22	EIF4B	HNRNPH2
Atx2	DAZAP1	EIF4E	HNRNPK
Atx2-L	DAZAP2	EIF4G1	HNRNPUL1
BAG3	DAZL	EIF4G2	HNRPA1
BANF1	DCD	EIF4H	HNRPK
BRAT1	DCP1A	ELAVL1	HNRPQ
BRF1	DCTN1	ELAVL2	HSF1
BRF2	DDX1	ELAVL4	HSP90AA1
C9ORF72	DDX19A	EPPK1	HSPA4
CALML5	DDX21	ETF1	HSPA9
CALR3	DDX3	EWS	HSPB1
CAP1	DDX47	FAK1	HSPD1
Caprin-1	DDX50	FAM120A	IFIH1
CAPZA2	DDX58	FAM195A	IGF2BP1
CARHSP1	DDX6	FAM195B	IGF2BP2
CASC3	DERA	FAM98A	IGF2BP3
CBFB	DHX30	FASTK	IP5K
CBX1	DHX36	FBL	IPO7
CCAR1	DISC1	FBP1	IPO8
CCDC9B	DKC1	FBP2	ITGB1

CCT3	DNAJA1	FHL1	KANK2
CCT6A	DNAJC8	FLNB	KHDRBS1
CD24	DNCI2	FMRP	KHDRBS3
CDC5L	DPYSL2	FNDC3B	KHSRP
CDC73	DPYSL3	FSCN1	KIF23
CDK1	DSP	FTSJ3	KIF5B
CDK2	DST	FUBP3	KLC1
CELF1	DSTN	FUS	KPNA1
CERKL	DTX3L	FXR1	KPNA2
CFL1	DYNC1H1	FXR2P	KPNA3
CHCHD3	DYNLL2	G3BP1	KPNA6
CHORDC1	DYRK3	G3BP2	KPNB1
L1RE1	NUFIP2	PTK2	SFRS3
LARP1	NUP205	PUM1	SGNP
LARP4	NUP98	PUM2	SIPA1L1
LBR	NXF1	PURA	SIRT6
LC3	OGFOD1	PURB	SLC6A8
LEMD3	OGG1	PXDNL	SMARCA1
Lin28A	OPTN	PYCR1	SMAUG
LMNA	PA2G4	QKI	SMC4
LPP	PABP1	RAB1A	SMG1
LSM14A	PABP4	RACGAP1	SMN
LSM14B	PABPC3	RACK1	SMU1
LSM3	PAK4	RAD21	SND1
LUC7L	PALLD	RANBP1	SNRPF
LUZP1	PARG	RAP55	SNTB2
MACF1	PARP1	RAPTOR	SORBS1
MAEL	PARP12	RBBP4	SORBS3
MAGEA4	PARP14	RBFOX1	SPAG5
MAGED1	PARP15	RBFOX2	SPATS2L
MAGED2	PAWR	RBM12B	SPECC1L
MAGOHB	PCBP2	RBM26	SQSTM1
MAP2K7	PCNA	RBM42	SRI
MAP4	PDCD6IP	RBMS1	SRP14
MAP4K4	PDLIM1	RBMS2	SRP9
MAPK8	PDLIM4	RBPMS	SRRT
MAPK8IP3	PDLIM5	RCC1	SRSF1
MAPRE1	PDS5B	RCC2	SRSF3
MARS	PELO	RED	SRSF4
MBNL	PFDN4	RENT2	STAT1
MCM4	PFN1	RFC3	STAU1
MCM5	PFN2	RFC4	STAU2
MCM7	PGAM5	RGPD3	STIP1
METAP1	PHB2	RhoA	STRAP
MEX3A	PHLDB2	RNF214	SUGP2
MEX3B	PKCα	RNH1	SUN1
MEX3C	PKP1	ROCK1	SYCP3
MFAP1	PKP2	ROQUIN	Syk
MKI67	PKP3	RPS19	SYNE1
MOV10	POLR2B	RPS2	TAF15
MSH6	PPIP5K1	RPS3	TCEA1
MSI1	PPME1	RPS3a	TCP1
MTHFSD	PPP1R10	RPS6	TDP-43

MSI2	PPP1R18	RPS6KA3	TDRD3
MTHFD1	PPP2R1A	RSL1D1	TERT
mTOR	PQBP1	PSMD2	TIA-1
MYO6	PRDX1	RTCB	TIAR
NCOA3	PRDX6	RTRAF	TMOD3
NEXN	PRKRA	S100A7A	TNKS
NONO	PRMT1	S100A9	TNKS1BP1
NOP58	PRMT5	SAFB2	TNPO1
NOSIP	PRRC2A	SEC24C	TNPO2
NSUN2	PRRC2C	SERBP1	TNRC6B
NTMT1	PTBP3	SFN	TOMM34
NUDC	PTGES3	SFPQ	TPM1
TPM2	TUBA4A	USP5	YWHAB
TPT1	TUBB3	VASP	YWHAH
TRAF2	TUBB8	VCP	YWHAQ
TRDMT1	TUFM	WDR62	ZBP1
TRIM21	TXN	XRN1	ZC3H12A
TRIM25	U2AF1	YARS	ZC3H14
TRIM56	UBA1	YBX1	ZC3H7A
TRIP6	UBAP2	YBX3	ZC3H7B
TSC1	UBAP2L	YES1	ZC3HAV1
TTP	UBL4A	YTHDF1	ZMAT3
TUBA1C	UPF1	YTHDF2	ZNF638
TUBA3C	USP10	YTHDF3	

Annex C. Cell types**.** Different cell lines used by the studies reporting the recruitment of the identified proteins to Stress Granules, listed in alphabetical order.

A172
A431
A549
BE-M17
BHK-21
BOSC
BxPc3
C2C12
C33A
CaCo2
CCL39
CHO-K1
Cortical neurons
COS-1
COS-7
DG75
DT40
DU145
EF

EL4
F11
F470
Fibroblast cell lines
H1299 (CRL-5803)
H35
H4
H9
HaCaT
HCT116
HDFs
HEC-1B
HEK293T
HeLa
Hep-2
HME
HPDE
HT1080
HUES1
HUES3
Human B-lymphoblastoid
Human SMA type I fibroblasts (3813 and 9677)
IMR-91
J1 ES
Jurkat
LCLs
LHCN-M2 human myoblasts
M57
MCF-10A
MCF7
MDA-MB231
MEFs
MiaPaCa-2
NIH 3T3
NRK
NSC-34
NT2
Oligodendrocytes
P19
PDAC
Rat hippocampal neurons
Rat oligodendrocytes
RAW 264.7
RDG3
RPE-1

Saos2
Schneider S2R+
SH-SY5Y
SK-N-BE(2)C
SKN-MC
STEK cells
SW13
SW480
T47D
THP-1
tsH1
U1242MG
U343
U-2 OS
VSMCS
W4 ES

Annex D. Stressors**.** Different stress stimuli that were used by the different studies reporting the recruitment of proteins to Stress Granules, listed in alphabetical order.

Arsenic oxide
Arsenic trioxide
Cisplatin
Clotrimazole
Cycloheximide
Dithiothreitol
Gamma rays (irradiation)
Heat shock
Hippuristanol (HIPP)
Hydrogen peroxide
Ionophore carbonyl cyanide m-chlorophenylhydrazone (CCCP)
Kinase inhibitors (Wortmannin, AMA-37, Ku55933, Rapamycin)
Leptomycin B
MG132
Nocodazole
Oxidant diethyl maleate (DEM)
Pateamine
PMA (phorbol 12-myristate 13-acetate)
Poly(I)(C) treatment
Puromycin
Sodium arsenite
Sodium pyruvate
Sodium selenite

Sorbitol
Taxol
Thapsigargin
Tunicamycin
UV irradiation