Chromatin-bound DYRK1A: promoter occupancy and implications in the regulation of ribosomal protein gene expression

Laura Barba Moreno

TESI DOCTORAL UPF / 2018

THESIS SUPERVISOR

Dra. Susana de la Luna

Gene Regulation, Stem Cells and Cancer Programme

Center for Genomic Regulation

DEPARTMENT OF EXPERIMENTAL AND HEALTH SCIENCE

A mi padre,

Table of contents

Results

Annexes

Abstract

DYRK1A (dual-specificity tyrosine-regulated kinase 1A) belongs to a conserved family of protein kinases present in all eukaryotes. DYRK1A is known to participate in cell proliferation and differentiation decisions and to fulfill key roles in brain development. Data from our laboratory indicate that DYRK1A acts as a transcriptional activator directly at proximal promoter regions of genes containing a conserved palindromic motif. In this Thesis work, the mechanism of DYRK1A recruitment to chromatin has been explored, as well as the DYRK1A transcriptional effect on a subset of gene targets, namely ribosomal protein genes (RPGs). By using electrophoretic mobility assays, DYRK1A has been found to interact directly with single-stranded DNA probes containing the consensus motif *in vitro*. Analysis of ChIP-Seq (chromatin immunoprecipitation coupled with high-throughput sequencing) data from ENCODE revealed that several transcription/chromatin-related factors are recruited to the DYRK1A-consensus motif. The list includes the co-repressor ZBTB33/KAISO and the tumor suppressor BRCA1. By means of genome-wide ChIP-Seq approaches, the presence of the two proteins has been established in DYRK1A-bound promoter regions enriched in the conserved motif; in fact, DYRK1A seems to be involved in BRCA1 recruitment to its genomic *loci,* probably in a catalytic-dependent manner since BRCA1 has shown to be a DYRK1A substrate. In addition, the analysis shows that the presence of DYRK1A at RPG promoters negatively correlated with that of the transcription factor GABP, suggesting the existence of two clusters of RPGs from the viewpoint of their transcriptional regulation. Indeed, the mRNA levels of chromatin-bound DYRK1A RPGs were significantly reduced in DYRK1A-silenced cells, which also showed an impairment in protein synthesis. Based on these results, DYRK1A might be a novel modulator of cell growth by regulating the expression of RPGs.

Resumen

DYRK1A (dual-specificity tyrosine-regulated kinase 1A) pertenece a una familia de proteína quinasas presente en todos los organismos eucariotas. DYRK1A participa en procesos de proliferación y diferenciación celular, y desempeña funciones clave durante el desarrollo del cerebro. Datos obtenidos en nuestro laboratorio indican que DYRK1A actúa como un activador transcripcional directamente a nivel de regiones promotoras de genes que mayoritariamente contienen una secuencia palindrómica conservada. En esta Tesis se han estudiado posibles mecanismos de reclutamiento de DYRK1A a cromatina, así como el efecto transcripcional de DYRK1A sobre un grupo de sus genes diana, en particular aquellos que codifican para proteínas ribosomales (GPRs). Mediante ensayos de cambio de movilidad electroforética, se ha demostrado que DYRK1A puede interaccionar *in vitro* directamente con sondas de ADN de hebra simple que contienen el motivo consenso. El análisis de datos de ChIP-Seq (inmunoprecipitación de cromatina asociada a secuenciación masiva) procedentes de ENCODE reveló que algunos factores relacionados con procesos de transcripción y regulación de la cromatina podrían ser reclutados al motivo asociado a DYRK1A. La lista incluye al co-represor ZBTB33/KAISO y al supresor de tumores BRCA1. Mediante estudios globales de ChIP-Seq se ha establecido el patrón de reclutamiento de las tres proteínas en la línea celular T98G, demostrando su presencia común en regiones promotoras enriquecidas en el motivo palindrómico. DYRK1A está involucrado en el reclutamiento de BRCA1 a sus correspondientes regiones genómicas, probablemente dependiendo de la actividad catalítica de la quinasa ya que los resultados indican que BRCA1 es un sustrato de DYRK1A. Adicionalmente, el análisis muestra que la presencia de DYRK1A en GPRs correlaciona negativamente con la del factor de transcripción GABP, sugiriendo la existencia de dos grupos de GPRs desde el punto de vista de regulación transcriptional. De hecho, los niveles de RNA mensajero de GPRs dianas de DYRK1A están significativamente reducidos en células donde la expresión de DYRK1A está silenciada; estas células también presentan una reducción de la síntesis proteica. Todos estos datos permiten proponer a DYRK1A como un nuevo regulador de la transcripción de GPRs, contribuyendo así al control del crecimiento celular.

Introduction

1. Dual-specificity tyrosine (Y)-regulated kinase 1A (DYRK1A)

1.1. DYRK subfamily of protein kinases

Protein kinases comprise a huge superfamily of proteins, which are classified according to the homology in their catalytic domain (Hanks and Hunter, 1995). Dual-specificity tyrosine (Y)-regulated kinases (DYRKs) belong to the CMGC group, also composed by Cyclin-dependent kinases (CDKs), Mitogen-activated protein kinases (MAPKs), Glycogen synthase kinases (GSKs), CDK-like kinases (CDKLs), Serine-arginine-rich protein kinases (SRPKs), Cdc2-like kinases (CLKs) and RCK family (Fig. I1A). DYRK family is, in turn, subdivided in DYRKs, Homeodomain-interacting kinases (HIPKs) and Pre-messenger RNA processing protein 4 kinases (PRP4Ks) subfamilies (Aranda et al., 2011) (Fig. I1B). Within the DYRK subfamily, three classes are found based on phylogenetic analysis: one class does not contain members in the animal kingdom and it is mainly represented by Yak1p from *Saccharomyces cerevisiae*, whereas class I and class II DYRKs comprise members from yeast to humans (Aranda et al., 2011) (examples found in Fig. I1B). From herein, DYRK will refer to the DYRK subfamily of kinases.

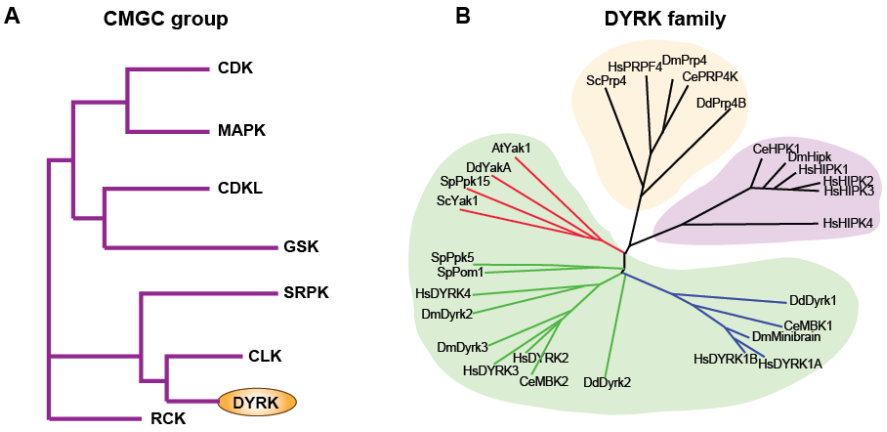


Figure I1: Phylogenetic classification of DYRK family of proteins. (A) Evolutionary relatedness among the families belonging to the CMGC group (see text for the complete name of the families). (**B)** Unrooted evolutionary tree depicting the DYRK subfamily members, PRP4K (in orange), HIPK (in violet) and DYRK (in green). At: *Arabidopsis thaliana*; Ce: *Caenorhabditis elegans;* Dd: *Dictyostelium discoideum;* Dm: *Drosophila melanogaster*; Hs: *Homo sapiens*; Sc: *Saccharomyces cerevisiae;* Sp: *Schizosaccharomyces pombe.* Adapted from Aranda et al., 2011.

In mammals, the phylogenetic and functional classifications are closely related (Han et al., 2012). Although the catalytic domain and the DYRK homology (DH)-box are highly conserved, class I and class II kinases

have different domains shown in Fig. I2 (Becker and Joost, 1999). Class II DYRKs are characterized by the presence of a N-terminal conserved domain known as N-terminal autophosphorylation accessory region (NAPA), which is essential for catalytic activation (Kinstrie et al., 2010) (Fig. I2).

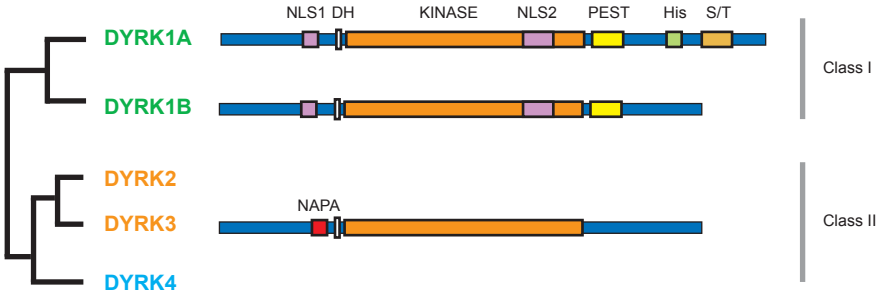


Figure I2: Classification and protein structure of the mammalian DYRK subfamily. Schematic representation of mammalian DYRK subfamily members indicating the protein domains: a common central kinase domain (KINASE) and the DYRK-homology box (DH). Class I presents two nuclear localization signals (NLS1 and NLS2) and a PEST region (PEST); DYRK1A has a stretch of 13 histidine residues (His) and a region enriched in serine and threonine residues (S/T). The non-catalytic NAPA domain, distinctive of class II DYRKs, is shown. The two classes in which DYRK members are divided are indicated on the right. Adapted from Aranda et al., 2011.

1.2. The protein kinase DYRK1A

DYRK1A was first discovered by inferring its sequence homology with the *mnb* gene in *D. melanogaster*. *DYRK1A* is located in the chromosome 21 in humans, within the Down syndrome (DS) critical region (DSCR) (Guimera et al., 1996). Actually, its huge impact in development and disease has prompted many scientists to carry out molecular studies, briefly summarized in this section.

1.2.1. Mechanism of activation

DYRK1A, like the other members of the DYRK family, is defined as a “dual-specificity” kinase because it is able to phosphorylate both tyrosine (Tyr) and serine/threonine (Ser/Thr) residues. Nevertheless, its capability of targeting Tyr is restricted only to autophosphorylation, being the phosphorylation of residue Tyr321, located within the activation loop, the essential event for the activation of the kinase (Himpel et al., 2001; Kentrup et al., 1996). Interestingly, neither upstream kinases targeting DYRK1A nor non-catalytic regions within the protein appear to mediate the activation process. Lochhead and colleagues proposed a mechanism by which an intermediate form of DYRK1A during translation would have different substrate specificity than the mature one thereby allowing Tyr autophosphorylation one time only (Lochhead et al., 2005).

Despite of being essential for activation, Tyr321 phosphorylation might not be required for the maintenance of activity, since dephosphorylation of this residue does not completely inactivate the kinase *in vitro* (Adayev et al., 2007; Becker and Sippl, 2011). Contrary to the generally accepted mechanism, some authors have claimed that DYRK1A is able to target Tyr in substrates and they argued that Tyr321 autophosphorylation still takes place in the mature form of DYRK1A (Walte et al., 2013).

Regarding phosphorylation site specificity, a consensus sequence has been proposed RPX(S/T)P (Himpel et al., 2000). The presence of a proline in the +1 position makes DYRK1A to be designated as a proline-directed kinase, although other small hydrophobic residues are allowed in this position (serine, alanine and valine) (Soundararajan et al., 2013).

Since DYRK1A is considered a constitutively active kinase, its function would be regulated at other levels, including protein expression, subcellular localization, protein stability and/or by undergoing interactor-directed conformational changes (detailed in sections I1.2.2-I1.2.5).

1.2.2. Gene and protein expression

DYRK1A is a ubiquitously expressed gene with a complex pattern of expression during development and in different tissues (Fig. I3A). Changes in RNA expression from embryos to adult are also observed in other tissues (mouse transcriptome data, Encyclopedia of DNA Elements [ENCODE]).

DYRK1A is expressed as two isoforms, different in 9 amino acids (aa), which are consequence of an alternative splicing event at the acceptor site of exon 4 (Guimera et al., 1999) (Fig. I3C). They are conserved in mammals and their expression levels are similar in cells and tissues (data not shown). In addition, the usage of alternative promoters (pM, pA and pB) might influence further 5'-untranslated region (UTR) splicing events resulting in a 29 aa shorter isoform when pB is used and skipping of exon 2 occurs (Maenz et al., 2008) (Fig. I3B). Several other alternative splicing isoforms affecting the C-terminus have been described (Guimera et al., 1996; Wang et al., 2017), although their expression is very low and their biological significance is unknown.

The promoters present a differential combination of transcription factor (TF) binding sites, pointing to differential regulation. A cAMP responsive element (CRE) is present within pA (Impey et al., 2004), whereas pB transcription activation seems to depend on E2F1 (Maenz et al., 2008). Moreover, an Activator protein 4 (AP4)-geminin repressor complex has been observed to bind pM and pB, allowing the recruitment of Silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) and

Histone deacetylase 3 (HDCA3), to attenuate DYRK1A expression during neuronal development (Kim et al., 2006b). Promoter pA also responds to RE1 silencing transcription factor (REST) via a neuron-restrictive silencer element (Lu et al., 2011), which might contribute to the transcriptional regulation of DYRK1A expression during neurodevelopment. In addition, the Nuclear factor of activated T-cells (NFAT) proteins or T-box transcription factor 5 have been described to induce DYRK1A during osteoclast differentiation and breast carcinogenesis, respectively (Kim et al., 2016; Lee et al., 2009).

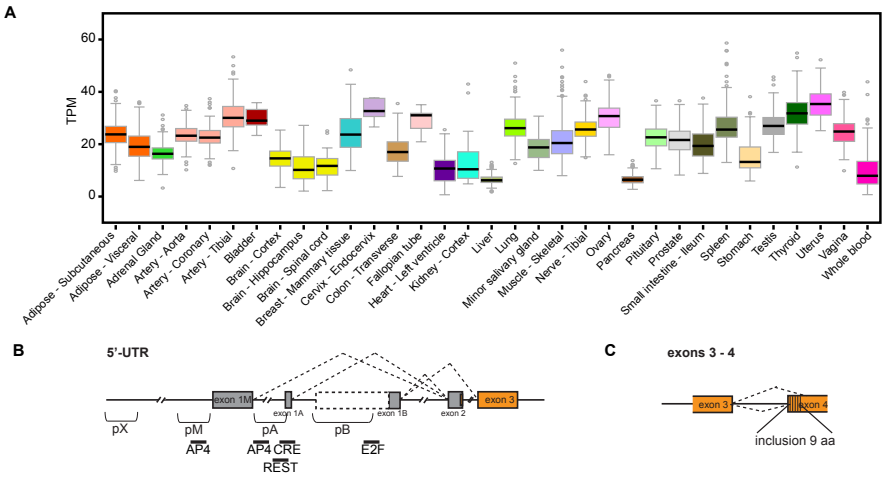


Figure I3: DYRK1A expression data across different human tissues. (A) Box plots showing the RNA expression levels of DYRK1A in TPM (transcripts per million) across different tissues. Data obtained from the Genotype-tissue expression (GTEx) portal. **(B, C)** Promoter usage and alternative spliced events in DYRK1A. See text for information on promoters pA, pB and pM. The existence of an upstream promoter (pX) is inferred from the presence of chromatin marks based on ENCODE data. Image adapted from Aranda et al., 2011.

On top of that, expression of DYRK1A is controlled post-transcriptionally. On one hand, micro-RNAs (miRNAs) such as miR-199b or miR-1246 have been proven to target DYRK1A (da Costa Martins et al., 2010; Zhang et al., 2011). On the other hand, regulatory mechanisms at the level of messenger RNA (mRNA) translation have been reported recently. Thus, during synaptogenesis, DYRK1A mRNA is directed for local translation at the axonal growth cones by a ribonucleoprotein complex containing the actin regulator Enabled homolog/Mammalian enabled (Vidaki et al., 2017).

1.2.3. Protein stability

DYRK1A is considered a short half-life protein due to the presence of a PEST domain, a region traditionally linked to rapid degradation (Rogers et al., 1986), in the non-catalytic C-terminus (Fig. I2). However, a clear

regulatory mechanism for how DYRK1A is targeted for degradation has not been elucidated yet. The protein is sensitive to proteasome inhibitors (Alvarez, 2004), indicating that its main route of degradation is the proteasome. However, no regulators of the process have been described yet, with a few exceptions like the E3 ligase $SCF^{\beta TrCP}$ or the chaperon Heat shock protein 90 (HSP90) (Liu et al., 2016; Sonamoto et al., 2015). In fact, DYRK1A has been shown to interact with HSP90 and the co-chaperone Cell division cycle 37 during its maturation process (Sonamoto et al., 2015). In addition, the catalytic activity could be an important requirement for stability since active DYRK1A has a longer half-life than the inactive form (Alvarez, 2004). In fact, some autophosphorylation sites such as Ser97 has been proven essential to increase stability (Kii et al., 2016). In this context, the Nemo-like kinase was shown to phosphorylate DYRK1A in several residues, and proposed to mediate its degradation via the proteasome pathway (Arato, 2010).

Furthermore, DYRK1A expression has been observed to be cell cycle-dependent in HeLa cells (Di Vona C, de la Luna S, unpublished results). Thus, in Hela cells protein levels increase in the S phase, in parallel with an increase in mRNA levels, and reach its maximum levels during G2/M, without changes in mRNA levels. The transcription-independent changes in protein amount could be due to protein stabilization/degradation (Di Vona C, de la Luna S, unpublished results).

1.2.4. Subcellular localization

DYRK1A presents a variable subcellular distribution depending on the cellular context. Thus, it has been found enriched both in the cytoplasm (i.e. glial cells in chick embryos; Hammerle et al., 2003) and in the nuclei of several cell types within the central nervous system (i.e. granular cell layer of the cerebellum [Marti et al., 2003], hippocampal neurons [Sitz et al., 2004], or embryonic neocortex neurons [Yabut et al., 2010]). These observations are in concordance with the identification of substrates in both compartments (Fig. I4). Particularly, the kinase has been found at specific cytosolic structures such as synaptic membranes, cytoskeleton and vesicle-containing fractions (Aranda et al., 2008; Kaczmarski et al., 2014; Murakami et al., 2009; Wegiel et al., 2004), as well as in nuclear speckles and at genomic regions (Alvarez et al., 2003; Di Vona et al., 2015; Jang et al., 2014; Salichs et al., 2009).

However, very little is known about the regulation of DYRK1A nucleocytoplasmic transport. Interestingly, exogenously expressed DYRK1A accumulates exclusively in the nucleus of several established cell lines (Alvarez et al., 2003; Becker et al., 1998; Sitz et al., 2004), leading to the belief that this may be the default localization when the

nucleus to cytoplasm traffic is saturated. Moreover, forced cytoplasmic accumulation is observed when the scaffold protein DNA damage-binding protein 1 and cullin 4-associated factor 7 and the adenovirus oncoprotein E1A are overexpressed (Glenewinkel et al., 2016), suggesting that some interactors might be needed to trigger the subcellular relocation.

Mechanistically, two NLSs have been characterized within the protein, a bipartite NLS in the non-catalytic N-terminus and a complex NLS in the CMGC insert within the catalytic domain (Fig. I2), and both have been proven necessary to direct DYRK1A to the nucleus (Alvarez et al., 2003). Moreover, the His tract localized in the C-terminal part of DYRK1A (Fig. I2) directs the protein to speckles, subnuclear organelles functioning as stores of factors implicated in transcription and pre-mRNA processing (Alvarez et al., 2003; Salichs et al., 2009). On top of that, different phosphorylated forms of DYRK1A have been identified differentially distributed across the subcellular compartments in human and mouse brain samples, pointing to post-translational modifications as an additional mechanism to regulate the subcellular distribution (Kaczmarski et al., 2014; Kida et al., 2011).

1.2.5. Modulators of the catalytic activity

Based on the activation mechanism (I1.2.1), it is reasonable to think that dephosphorylation of the activation loop would lead to the inactivation of the kinase, at least partially. However, no DYRK1A phosphatases have been identified so far. Nevertheless, some proteins have been identified as modulators of the kinase activity.

The scaffold proteins 14-3-3 enhance DYRK1A intrinsic catalytic activity by binding two DYRK1A sites: one located at the N-terminal domain, and other within the PEST domain, containing a Ser520 residue that needs to be autophosphorylated to be recognized by 14-3-3β (Alvarez et al., 2007; Kim et al., 2004). Another putative activator is Large tumor suppressor kinase 2, which phosphorylates DYRK1A and promotes DYRK1A-dependent phosphorylation of Lin52, a component of the Dimerization partner, retinoblastoma [Rb]-like, E2F and multi-vulval class B (DREAM) complex (Tschop et al., 2011). In addition, Sprouty-related-EVH1 domain-containing protein 1/2 (SPRED 1/2) have been shown to hinder accessibility to DYRK1A substrates (Li et al., 2010).

1.3. DYRK1A and disease

As previously mentioned, *DYRK1A* is located within the DSCR, and it is upregulated 1.5-fold in DS individuals (Dowjat et al., 2007; Guimera et al., 1996). Mouse models overexpressing *DYRK1A* reproduce a wide range of

the features associated to DS: intellectual disability (Altafaj et al., 2001; Dowjat et al., 2007), motor alterations (Altafaj et al., 2001; Martinez de Lagran et al., 2004), retinal abnormalities (Laguna et al., 2008), skeletal alterations (Blazek et al., 2015; Lee et al., 2009) or higher risk to suffer childhood leukemia (Malinge et al., 2012). In fact, it has been demonstrated that the cognitive and motor defects, the retinal abnormalities, the Alzheimer-like and the skeletal phenotypes are attenuated when DYRK1A levels are normalized, either by genetic or pharmacological means (Blazek et al., 2015; Garcia-Cerro et al., 2014; Garcia-Cerro et al., 2017; Laguna et al., 2013; Ortiz-Abalia et al., 2008).

DYRK1A is highly sensitive to gene dosage imbalance: not only when in trisomy but when in haploinsufficiency it is linked to clinical phenotypes. The list of mutations of diverse nature (translocations, insertions/deletions, single nucleotide variations, etc.) affecting one *DYRK1A* allele has increased over the last years. Clinical traits include developmental delay, microcephaly, seizures, feeding problems, a characteristic facial gestalt and speech impairment (Luco et al., 2016 and references therein). Currently, *DYRK1A* haploinsufficiency is considered a rare clinical syndrome, within the autism spectrum disorders (ASD), in the Online Mendelian Inheritance in man database (OMIM: 614104) and in Orphanet (ORPHA: 464306) (see www.ncbi.nlm.nih.gov/books/NBK333438 for clinical description and management). In agreement, a mouse model heterozygous for *Dyrk1a* present low neonatal viability, developmental delay, small body size, microcephaly, seizures and behavioral traits of ASD (Arranz, 2016; Fotaki et al., 2002), and a mouse model recreating one of the *DYRK1A* frameshift mutations suffers cognitive impairments (Raveau et al., 2018).

In addition, and given that DYRK1A phosphorylates key factors in the development of neurodegenerative diseases such as Amyloid precursor protein (APP), Tau, α-Synuclein and Huntingtin interacting protein 1 (HIP1) (Kang et al., 2005; Kim et al., 2006a; Ryoo et al., 2008; Ryoo et al., 2007), dysregulation of DYRK1A has been linked to the development of Alzheimer's, Parkinson's and Huntington's disease (Barallobre et al., 2014; Branca et al., 2017; Cen et al., 2016; Kimura et al., 2007).

Furthermore, DYRK1A dysregulation might lead to heart defects due to alterations in the Rb-E2F and NFAT pathways (da Costa Martins et al., 2010; Hille et al., 2016; Kuhn et al., 2009; Raaf et al., 2010), and it has been also related to osteoporotic phenotypes (Lee et al., 2009). Apart from that, DYRK1A inhibition has been considered a potential therapeutic target to stimulate pancreatic β-cells proliferation via NFAT signaling, thereby ameliorating diabetes (Belgardt and Lammert, 2016). Finally, a

potential link of DYRK1A with cancer is suspected, and a few reports on this line have appeared in recent years (Liu et al., 2014; Pozo et al., 2013; Radhakrishnan et al., 2016), with DYRK1A playing positive and negative roles depending on different tumor types. Indeed, links with cell proliferation and differentiation have been established in different cellular contexts. On one hand, DYRK1A is a negative regulator of cell cycle since it targets Cyclin D1 for degradation, it is able to increase the levels of the CDK inhibitor p27, and favor DREAM complex assembly allowing cells to enter in a quiescent state (Chen et al., 2013; Litovchick et al., 2011; Soppa et al., 2014). On the other hand, it could promote proliferation by facilitating nuclear translocation of Glioma-associated oncogene homolog 1 (GLI1) (Mao et al., 2002) or by stimulating the Epidermal growth factor receptor pathway (Pozo et al., 2013) in some tumor contexts. Finally, DYRK1A potentiates cell survival by targeting Caspase 9 (Laguna et al., 2008) or Sirtuin 1, which lately deacetylates p53, thus skipping apoptosis (Guo et al., 2010).

1.4. DYRK1A inhibitors

The impact of DYRK1A overexpression in health aimed to investigate compounds that could inhibit the activity of the kinase. The most used one in research is the adenosine triphosphate (ATP) competitor harmine, a β-carboline alkaloid able to target Tyr321 autophosphorylation during the maturation process (Bain et al., 2007; Gockler et al., 2009). Despite attenuating some features attributed to DYRK1A overexpression in mouse models (Kuhn et al., 2009; Laguna et al., 2008; Pozo et al., 2013), its capability to inhibit the Monoamine oxidase enzyme difficulties its entry into clinical studies. A polyphenol derivative found in the green tea leaves, Epigallocatechin gallate (EGCG), with anti-tumor activities (Yang et al., 2009), has proven effective as a DYRK1A inhibitor both *in vitro* and *in vivo* (De la Torre et al., 2014; Guedj et al., 2009; McElyea et al., 2016). In fact, EGCG is under clinical trial to treat DS individuals (de la Torre et al., 2016). Nevertheless, both compounds target other members of the DYRK subfamily, which makes the search for alternative drugs necessary.

In the last years, derivatives of indirubin (Myrianthopoulos et al., 2013), roscovitine (Demange et al., 2013), leucettine (Tahtouh et al., 2012) and benzothiazole have been developed. Belonging to the last category, INDY and FINDY appear to be promising since they attenuate head malformations induced by overexpressed DYRK1A in *Xenopus laevis* (Kii et al., 2016; Ogawa et al., 2010). Other type of drugs, like Silmitasertib, a Casein kinase 2 (CK2) inhibitor, rescue neurological and phenotypic defects in an mnb-overexpressing *Drosophila* model (Kim et al., 2016). However, the lack of a DYRK1A-specific inhibitor persists. Recently, drugs

deriving from carbonitrile appeared to be highly selective for DYRK1A *in vitro,* although its poor aqueous solubility needs to be improved for a better bioavailability *in vivo* (Meine et al., 2018).

1.5. DYRK1A as a pleiotropic regulator of cellular processes

DYRK1A has been associated to a wide range of substrates both in the cytoplasm and in the nucleus (Fig. I4), which would be in agreement with its ability to be present in both compartments.

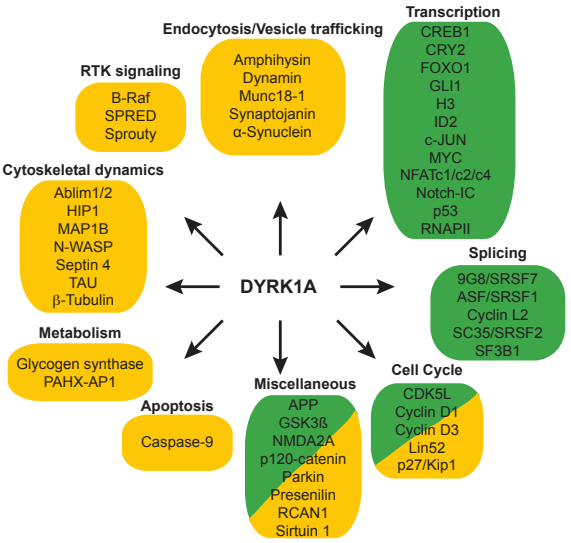


Figure I4: Substrates for mammalian DYRK1A. DYRK1A substrates are grouped according to their associated function within the cell, indicated on top of the box. Yellow boxes indicate proteins found in the cytosol or exerting cytosolic activities; green boxes indicate proteins with nuclear activities. For complete names, see abbreviations section.

Since the mechanism of nucleocytoplasmic trafficking is not fully understood yet, and most of its substrates are shuttling proteins, it is difficult to establish where the phosphorylation events take place. Nevertheless, the diverse nature of its substrates indicates that DYRK1A is a pleotropic protein involved in different processes like cell proliferation, cytoskeletal dynamics or vesicle trafficking, among others (Fig. I4). Of note, most of the phosphorylation events described to date have been shown only *in vitro*, thus the link of DYRK1A to some of the cellular processes associated to the substrates needs to be further explored. In addition, with the available data it is difficult to establish whether DYRK1A is the main effector or if it is playing a modulatory role in some processes. In fact, DYRK1A is a priming kinase for GSK3, facilitating GSK3 phosphorylation of Tau, Microtubule-associated protein 1B (MAP1B) or Cryptochrome circadian clock 2 (CRY2) (Kurabayashi et al., 2010; Scales et al., 2009; Woods et al., 2001), indicating that some of DYRK1A cellular activities could require actions in concert with GSK3.

As shown in Fig. I4, one of the functional categories is transcription, including well-known transcription factors such as GLI1 (Mao et al., 2002) or NFAT (Arron et al., 2006; Gwack et al., 2006). In addition, work from our laboratory demonstrated that DYRK1A affects transcription by being recruited directly on a subset of promoters (Di Vona et al., 2015). This role of DYRK1A will be detailed in the next section.

2. DYRK1A and transcription

2.1. Principles of eukaryotic RNA polymerase II-mediated transcription

Transcription is an essential step in gene expression regulation. In higher eukaryotes, there are three RNA polymerases that, despite sharing some subunits, interact with different combinations of TFs to precisely transcribe specific sets of nuclear genes (Zawel and Reinberg, 1995). RNA polymerase I (RNAPI) is dedicated to the transcription of the RNA precursor deriving in 18S, 5.8S and 28S ribosomal RNAs (rRNAs) (Grummt and Langst, 2013); RNA polymerase II (RNAPII) transcribes mRNAs that would be further translated into proteins and a wide variety of non-coding RNAs such as long non-coding RNAs and miRNAs; finally, RNA polymerase III (RNAPIII) is responsible for the transcription of small non-coding RNAs such as 5S rRNA or transfer RNAs (tRNAs) (Turowski and Tollervey, 2016). For the purpose of the thesis work, this section is mainly focused on basic concepts regarding RNAPII-mediated transcription, highlighting the initiation phase.

The first step is the assembly of the pre-initiation complex (PIC), which requires the 12 core subunits forming RNAPII and general TFs (GTFs), including TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. TFIID is composed by the TATA-box-binding protein (TBP) plus around 13-15 TBP-associated factors (TAFs), and it is responsible for the first contacts with core promoter elements. Thus, TBP recognizes the TATA-box while TAF6 and TAF9 binds the Downstream promoter element (DPE); TAF1 and TAF2 interact with the Initiator element (Inr) and TAF1 recognizes the Downstream core element (DCE). Later on, during PIC assembly, TFIIB will detect the TFIIB recognition element (BRE) (Maston et al., 2006; Vo Ngoc et al., 2017). Finally, the PIC is stabilized and directed to the transcriptional start site (TSS) with the participation of additional activators and co-activators, which interact with either their corresponding motifs along the proximal promoter region or via other previously recruited proteins, such as Mediator (Maston et al., 2006; Soutourina, 2017) (Fig. I5A).

One key player in the regulation of RNAPII transcription is the carboxy-terminal domain (CTD) of the subunit 1 of the holoenzyme. The CTD consists of multiples repeats (52 in mammals) of the heptapeptide Tyr-Ser-Pro-Thr-Ser-Pro-Ser, which are modified post-translationally by a plethora of proteins to achieve a tight regulation, not only of the transcription itself, but also to correctly orchestrate transcription with co-transcriptional processes such as mRNA capping or splicing (Harlen and Churchman, 2017). After some abortive rounds, RNAPII is able to escape the promoter thanks to stabilizing signals such as the phosphorylation of the CTD on Ser5 and Ser7 by the TFIIH complex formed by the CDK7 and cyclin H (Harlen and Churchman, 2017) (Fig. I5A).

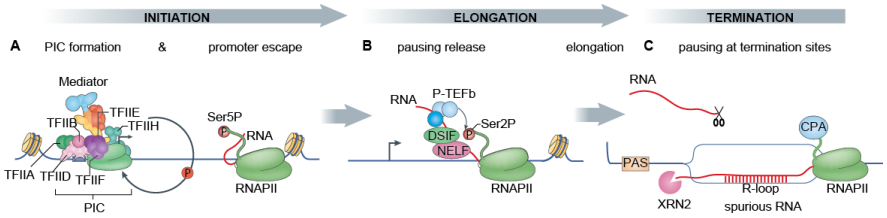


Figure I5: RNAPII-mediated transcription. Cartoon depicting the different phases of RNAPII transcription: initiation, elongation and termination and the main components involved. **(A)** PIC assembly and promoter escape, **(B)** pausing release, and **(C)** pausing at termination sites. See the text for details. Adapted from Skalska et al., 2017 and Soutourina, 2017.

Right after elongations starts, RNAPII enters in a pausing status mediated by Negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF) interaction. As a regulatory mechanism, RNAPII pause release only occurs when the positive transcription elongation factor (pTEFb) complex (formed by CDK9 and alternative Cyclins T1/T2a/T2b/K) phosphorylates both factors and the CTD at Ser2, thereby allowing productive elongation to take place (Saunders et al., 2006) (Fig. I5B). Finally, RNAPII slows down once it reaches the termination sequence known as polyA-site (PAS) likely due to: i) the recruitment to the CTD of an essential complex for final mRNA processing, 3'-end cleavage and polyadenylation complex (CPA), and ii) the formation of R-loops, namely, hybrids between the nascent RNA and the DNA template. CPA releases the transcript whereas RNAPII continues transcribing until the exoribonuclease XRN2, while degrading this spurious product, induces a conformational change in RNAPII that releases it from the DNA template (Proudfoot, 2016) (Fig. I5C).

2.2. DYRK1A as a direct regulator of transcription

For decades, kinases were thought to play an indirect role on transcription by targeting TFs or chromatin modifiers. However, it has become more and more evident that kinases phosphorylate proteins involved in chromatin functions (histones, chromatin remodelers, RNA polymerase components, etc.) *in situ*. Well-known examples belong to the MAPK family. Thus, c-Jun N-terminal kinase is recruited to active promoters during neurodifferentiation to phosphorylate histone 3 (H3) in residue Ser10 (Tiwari et al., 2011), Extracellular signal-regulated kinase 2 (ERK2) was shown to directly bind a specific DNA sequence contained in some interferon-induced gene promoters to exert a repressive function on their transcription (Hu et al., 2009), and p38α targets genes during skeletal muscle cell differentiation (Segales et al., 2016).

In addition, it has been shown that cells use kinases to phosphorylate the CTD other than CDK7 or CDK9. Thus, ERK2 targets Ser5 in the presence of Polycomb repressive complex 2 in developmental genes (Tee et al., 2014), whereas Ser2 phosphorylation can be carried out by CDK12, CDK13 or Bromodomain containing protein 4 (Jeronimo et al., 2016). In fact, CDK12 is thought to be the main kinase phosphorylating this residue in *Drosophila* (Bartkowiak et al., 2010).

The first indication of DYRK1A as a putative direct regulator of transcription appeared in (Himpel et al., 2000), where H3 was found to be phosphorylated by DYRK1A at Thr45 *in vitro.* Later on, DYRK1A was observed to transactivate steroid receptor responsive genes by acting synergistically with Androgen receptor interacting protein 4 in a kinase-independent manner (Sitz et al., 2004). Finally, DYRK1A was shown to phosphorylate *in vivo* H3 at Thr45 and Ser57 affecting the repressor function of Heterochromatin protein 1 on cytokine-respondent genes (Jang et al., 2014).

The work of our group showed that DYRK1A is seated on a subset of RNAPII proximal promoter regions based on chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) studies. The study also showed that DYRK1A phosphorylates the CTD on Ser2 and Ser5 *in vitro* and *in vivo*, leading to the proposal that this is the mechanism to regulate the transcriptional activation of its targets (Fig. I6) (Di Vona et al., 2015). Moreover, DYRK1A was linked to RNAPIII transcription since the second most represented category of DYRK1A-associated genomic *loci* correspond to intergenic regions (Fig. I6B), mostly tRNA genes, and some of them showed reduced transcription when DYRK1A expression was reduced (Di Vona, 2013).

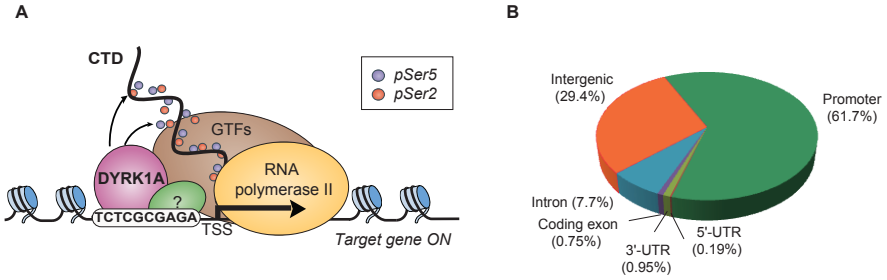


Figure I6: DYRK1A is recruited to proximal promoter regions. (A) Cartoon showing the working model proposed for DYRK1A-dependent transcription activation of gene targets. See text for details. GTFs are depicted as well as unknown putative DYRK1A binding partners in this context **(B)** Pie chart representing the distribution of DYRK1A genomic *loci* over the distinct genomic features in T98G cells as appeared in Di Vona et al., 2015.

DYRK1A-associated RNAPII promoters are enriched in a palindromic sequence (TCTCGCGAGA, herein named as “DYRK1A-motif") (Fig. I6A), which was proven to be important for the DYRK1A-dependent transcriptional activation function (Di Vona et al., 2015). Electrophoretic mobility shift assays (EMSAs) showed that this particular motif could serve as a platform for the recruitment of protein complexes where DYRK1A was found (Di Vona et al., 2015). Results from *de novo* discovery of motifs within the ENCODE Consortium project identified some factors whose associated-genomic *loci* were enriched in this particular motif: the transcriptional co-repressor Zinc finger and BTB domain containing 33 (ZBTB33/KAISO; from herein KAISO), the E3 ubiquitin ligase Breast cancer type 1 susceptibility protein (BRCA1) and the chromatin remodeler Chromodomain helicase DNA binding protein 2 (CHD2) (Wang et al., 2014). Moreover, further analysis of data from the ENCODE project uncovered additional factors associated to TCTCGCGAGA-enriched DNA sequences: the E26-transformation specific (ETS) family and Nuclear receptor subfamily 3 group C member 1 protein (NR3C1) (Kheradpour and Kellis, 2014). However, over-representation of the motif only appeared in one NR3C1 dataset, which was considered unreliable by the authors. For the ETS family, it was not known which member or members of this huge family of TFs were included in the study. For these reasons, the interest of this Thesis work on the question of whether there is a cross-talk between DYRK1A and other motif-binders was limited to CHD2, KAISO and BRCA1. A brief introduction of these three proteins is provided below.

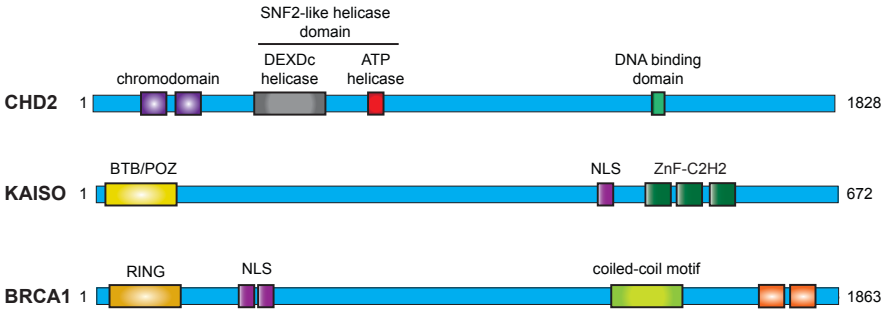


Figure I7: Schematic representation of the protein structure of CHD2, KAISO and BRCA1 proteins. See text for details. Drawings are not to scale.

2.3. Putative DYRK1A binding partners at the chromatin level

2.3.1. Chromodomain helicase DNA binding protein 2 (CHD2)

CHD2 is a chromatin modifier belonging to the evolutionarily conserved superfamily of Sucrose non-fermentable 2 (SNF2)-related ATPases, which change histones-DNA contacts in an ATP-dependent manner, and alter the structure of nucleosomes affecting chromatin accessibility (Nagarajan et al., 2009). In addition to the SNF2-like helicase domain, CHD subfamily members are characterized by the presence of N-terminally extra tandems of chromatin organization modifier (chromo) domains, which are in charge of interacting with heterochromatin. As a member of class I CHDs, CHD2 contains a DNA binding domain (DBD) in the C-terminus recognizing AT-rich DNA motifs (Fig. I7A).

CHD2 depletion causes reduced histone variant H3.3 occupancy on developmental regulated genes (Semba et al., 2017; Siggens et al., 2015). Mutations in *CHD2* have been linked to defects in brain development, leading to epilepsy and intellectual disability (Chenier et al., 2014). Moreover, *CHD2* is one of the most frequently mutated genes in chronic lymphocytic leukemia patients. In fact, *Chd2* heterozygous mice present increased hematopoiesis and susceptibility to lymphomas (Nagarajan et al., 2009; Rodriguez et al., 2015). In this context, CHD2 has been proposed to act as a modulator of DNA damage responses (DDR) (Nagarajan et al., 2009).

2.3.2. Zinc finger and BTB domain containing 33 (ZBTB33/KAISO)

KAISO was classified as a methyl DNA-binding protein, generally known as a transcriptional repressor. KAISO primary structure includes three zinc finger (ZF) domains located in the C-terminus and acting as DBDs (Fig. I7B), which recognize, on one hand, symmetrically methylated-cytosine/guanine dinucleotides (mCpGs) and, on the other hand, an

unmethylated CTGCNA motif called KAISO binding site (KBS) (Daniel, 2007; Filion et al., 2006) (Fig. I8). KAISO interacts with many diverse proteins through an N-terminal Broad complex, tramtrack, bric à brac/Pox virus and zinc finger (BTB/POZ) domain (Fig. I7B and I8). Thus, it forms homodimers and interacts with Nuclear receptor co-repressor 1 (NCoR1) to recruit histone deacetylases to mCpG promoter regions of the *Metastasis associated protein 2 (MTA2) locus* (Yoon et al., 2003) (Fig. I8A). KAISO also binds the enhancer blocker CCCTC-binding factor (CTCF), and negatively regulates its activity possibly by recognizing a KBS close to CTCF binding sites in the case of the human 5' β-globin insulator (Defossez et al., 2005). Cooperation between KAISO and CTCF has also been shown for the *RB* gene promoter and for Imprinting control region I (ICRI) (Bohne et al., 2016; De La Rosa-Velazquez et al., 2007) (Fig. I8B).

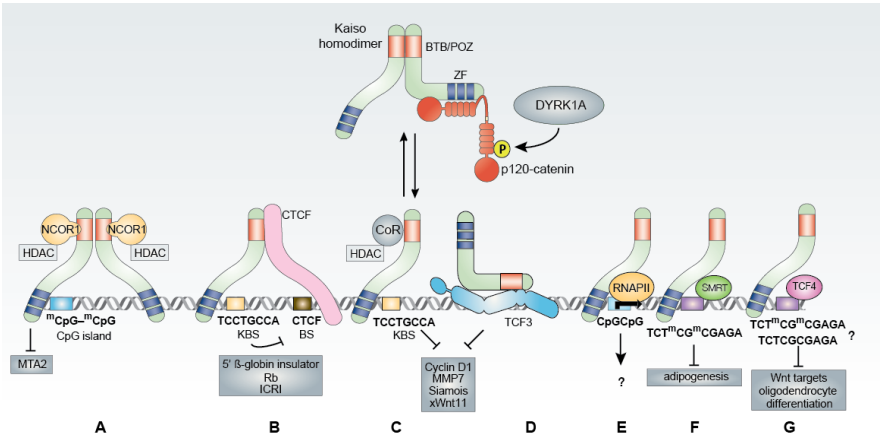


Figure I8: KAISO recruitment to DNA. (A) KAISO binds methylated CpGs and transcriptional repression of *MTA locus* is achieved via recruitment of co-repressors (COR/NCOR) and HDACs. **(B)** CTCF activity is blocked by KAISO recruitment to a KBS close to CTCF binding sites (BS). **(C)** KAISO binds DNA directly through KBSs and this process is counteracted when KAISO homodimers are sequestered by p120-catenin, which is promoted by DYRK1A phosphorylation. **(D)** KAISO interferes with canonical Wnt gene activation by interacting with TCF3. **(E)** KAISO is recruited to unmethylated CpGs enriched in RNAPII likely to activate transcription. **(F)** KAISO interacts with the TCTCGCGAGA-motif when methylated to inhibit the transcription of genes involved in adipogenesis in collaboration with SMRT. **(G)** KAISO interacts with the DYRK1A-motif to regulate genes involved in oligodendrocyte maturation together with TCF4. Adapted from van Roy and McCrea, 2005.

KAISO targets are closely connected to the Wnt signaling pathway, mainly involved in development and tumor progression processes. Upon Wnt signals, p120-catenin binds the ZF domains of KAISO, forcing its release from Wnt targets containing KBSs, and thereby inducing the expression of Wnt targets (Daniel and Reynolds, 1999; Kim et al., 2002). Interestingly,

DYRK1A favors this interaction by phosphorylating p120-catenin in the context of *X. laevis* embryo gastrulation (Hong et al., 2012). In addition, KAISO has been observed to play a role in the canonical Wnt pathway by cooperating with T-cell factor 3 (TCF3) to silence Wnt/β-catenin target genes (Park et al., 2005) (Fig I8C-D).

KAISO has not only been linked to transcriptional repression but also to activation. A genome-wide study revealed that KAISO binds mostly unmethylated RNAPII-occupied promoter regions decorated with histone marks associated to transcription activation (Blattler et al., 2013) (Fig. I8E). In agreement, KAISO has been shown to activate the transcription of Cyclins D1 and E1, leading to proliferation in HeLa cells (Pozner et al., 2016).

As mentioned before, KAISO has been associated with the DYRK1A-motif sequence in the context of the ENCODE project (Wang et al., 2014). KAISO's association with this motif is linked to a cooperative role with SMRT in adipogenesis and with TCF4 in oligodendrocyte differentiation (Raghav et al., 2012; Zhao et al., 2016) (Fig. I8F-G).

2.3.3. Breast cancer type 1 susceptibility protein (BRCA1)

BRCA1 was first discovered by looking for genes whose mutation led to hereditary early-onset breast cancer (Hall et al., 1990). *BRCA1* mutations display pleotropic effects, leading mainly to genetic instability so that BRCA1 is considered a tumor suppressor. Functionally, BRCA1 acts as an E3 ubiquitin ligase, in collaboration with BRCA1-associated Really interesting new gene [RING] domain protein 1 (BARD1), which depends on an N-terminal RING finger domain conferring the activity while a BRCA1 C-terminal (BRCT) domain is responsible for the majority of the protein-protein interactions (Roy et al., 2011) (Fig. I7C).

BRCA1 is mostly known by its role in DNA repair of double-strand breaks (DSBs), since it triggers recombinase RAD51 location to the damaged DNA so that homologous recombination repair can take place (Roy et al., 2011; Wu et al., 2010) (Fig. I9A and B). In addition, BRCA1 also participates in the checkpoints controlling the DNA status in the different phases of the cell cycle. BRCA1 is phosphorylated by Ataxia-telangiectasia mutated (ATM) and Ataxia-telangiectasia Rad3-related (ATR) during the DDR (Cortez et al., 1999; Tibbetts et al., 2000), and it is also phosphorylated by other kinases (Fig. I9A). The kinases include CDK1 or CDK2 during the DNA synthesis checkpoint (Johnson et al., 2009; Ruffner et al., 1999), Check point kinase 2 (CHK2) or Aurora-A during the G2/M phase (Deng, 2006; Zhang et al., 2004) and Polo-like kinase 1 (PLK1) in the DDR (Chabalier-Taste et al., 2016). In addition,

Fyn-related kinase (FRK/RAK), and Protein kinase B (PKB/AKT) also target BRCA1 to regulate its protein stability and nuclear localization, respectively (Hinton et al., 2007; Kim et al., 2017). The fact that several kinases target multiple residues of BRCA1 points to a highly complex regulatory network to accomplish BRCA1 coordinated activity.

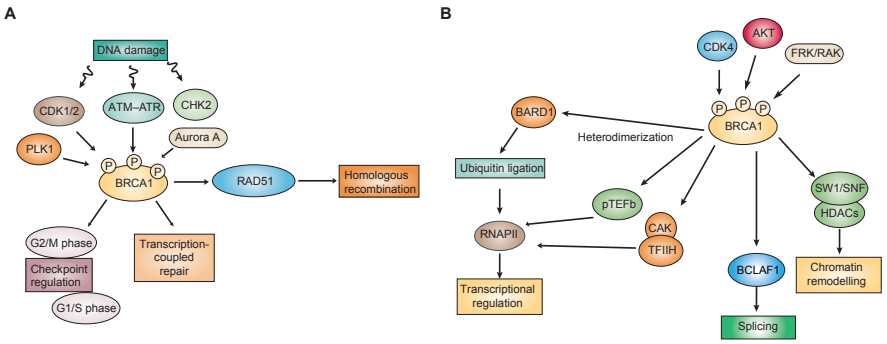


Figure I9: BRCA1 is a pleiotropic protein. Cartoon depicting the main pathways in which BRCA1 is involved. **(A)** DNA damage is a key trigger of BRCA1 activation. Damage sensors kinases (CDK1/2, ATM, ATR, CHK2, PLK1 and Aurora A) phosphorylate BRCA1 leading to BRCA1-dependent activation of DDR, which includes DNA damage repair via homologous recombination or coupled to transcription as well as cell cycle checkpoint activation in G1/S and G2/M. **(B)** Additional BRCA1 functions: see text for description. Upstream kinases triggering some of these responses are CDK4, AKT or FRK/RAK. Adapted from Narod and Foulkes, 2004.

Several studies have shown the association between BRCA1 and the RNAPII holoenzyme as well as its influence on the mRNA levels of some DDR genes, pointing to a possible role of BRCA1 in transcription (Nadeau et al., 2000). Interestingly, BRCA1 has been proposed as a CTD modulator, since it abrogates Ser5 phosphorylation by inhibiting the activity of CDK-activating kinase (CAK) when being part of a TFIIH sub-complex, or Ser2 phosphorylation by affecting the function of pTEFb (Moisan and Gaudreau, 2006; Moisan et al., 2004) (Fig. I9B). In addition, BRCA1 has been suggested to participate in the transcription-coupled repair process, a nucleotide excision repair pathway activated when DNA damage occurs during transcription (Fig. I9A). Thus, BRCA1 has been observed to accumulate in RNAPII pausing sites, usually characterized by the formation of R-loops, to mediate in the recruitment of the repair machinery (Hatchi et al., 2015; Zhang et al., 2017), and to target elongating RNAPII for degradation (Kleiman et al., 2005) (Fig. I9B). Other transcription-associated processes require the participation of BRCA1 in chromatin remodeling via interacting with factors such as Switch/SNF (SWI/SNF) complex (Bochar et al., 2000; Harte et al., 2010) or HDAC1 and HDAC2 (Yarden and Brody, 1999) (Fig. I9B). Finally, BRCA1 contributes to further gene regulation processes such as splicing, by taking part of a complex together with B-cell lymphoma 2-associated

factor 1 (BCLAF1) to mediate the splicing of pre-mRNAs whose proteins are involved in the DDR (Savage et al., 2014).

2.4. DYRK1A genomic targets connects DYRK1A with ribosome-related processes

Gene ontology analysis performed on DYRK1A-associated genomic *loci* revealed that DYRK1A mostly occupies promoter genes related to translation and ribosome biogenesis processes (Di Vona et al., 2015). Among these DYRK1A targets, there is a clear enrichment of genes encoding for canonical ribosomal proteins (RPs). In the next part of the Introduction, I will give an overview of what is currently known about RP gene (RPG) expression regulation, highlighting its main physiological implications in translation, cell growth and disease.

3. Ribosomal proteins transcriptional regulation

3.1. Principles of ribosome biogenesis and translation machinery

Ribosomes are huge cellular machineries in charge of the essential role of translating mRNAs into proteins. In mammals, they are formed by two subunits: the large 60S subunit is composed by the 5S, 5.8S and 28S rRNAs plus 52 RPs (RPLs), while the small 40S subunit comprises the 18S rRNA plus 35 RPs (RPSs). Recently, a new system for naming RPs has been proposed (Ban et al., 2014) (equivalences to the old nomenclature are listed in Annex I), and both nomenclatures will be used unless otherwise specified.

Ribosome biogenesis is complex and energy-consuming: it takes place in three different subcellular compartments (nucleolus, nucleoplasm and cytoplasm) and involves more than 200 factors of different nature (rRNAs, small nucleolar RNAs [snoRNAs] and canonical and auxiliary proteins) (Pena et al., 2017). The three RNA polymerases participate in the transcription of the ribosomal components. The rRNAs 18S, 5.8S and 28S are generated from the processing of a 47S pre-rRNA transcribed by RNAPI in the nucleolus, whereas the 5S rRNA is produced by RNAPIII in the nucleoplasm. In addition, all the coding genes involved in ribosome biogenesis are transcribed by RNAPII; their mRNAs are exported to the cytoplasm to be translated and once the proteins are produced, they must be transported into the nucleolus, not only to be part of the ribosome assembly, but also to help in the processing of rRNA (Pena et al., 2017; Xue and Barna, 2012) (Fig. I10). Finally, the two ribosomal subunits are

exported to the cytosol, where they assemble together and, after the release of assembly and transport factors, they are competent for translation.

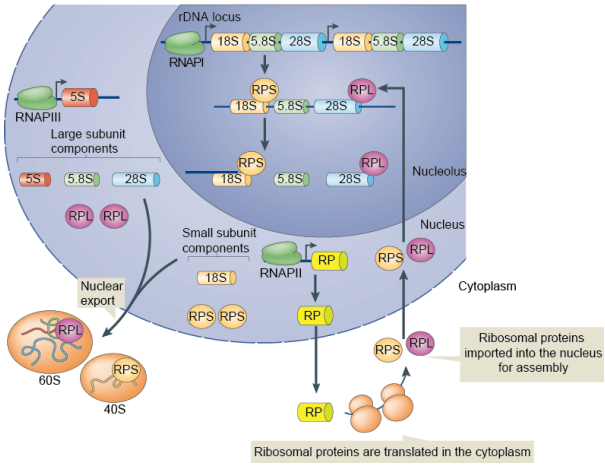


Figure I10: Schematic representation of the process of ribosome formation. See main text for details. Adapted from Xue and Barna, 2012.

Translation is organized in three steps, initiation, elongation and termination and requires, not only the ribosomal subunits, but also a set of distinct factors involved in each of the steps (Andreev et al., 2017; Dever and Green, 2012; Hinnebusch, 2014). All the three steps are subject of regulatory mechanism, which are the end point of cellular and environmental signals (Richter and Coller, 2015).

Translation initiation begins when, on one hand, the 43S PIC is assembled, and on the other hand, a post-transcriptional element present in the majority of mRNAs, m^7G cap, is recognized by the trimeric initiation complex eIF4E/G/A resulting in the formation of the 48S complex. Once the first AUG codon is recognized, the complex stops scanning the mRNA, the large subunit is recruited and several of the initiation factors are released (Andreev et al., 2017; Hinnebusch, 2014) (Fig. I11).

The new-formed 80S initiation complex (IC) continues translation elongation where aa-tRNA anticodons base-pair with their corresponding mRNA triplets. These events trigger a conformational shift of the ribosome allowing the formation of the peptide bond between the two aa and the displacement of the ribosome across the mRNA. Thereby, the former aa-tRNA can leave whereas the latter tRNA carries the peptide chain in formation (Dever and Green, 2012; Richter and Coller, 2015). Translation termination occurs when the ribosome meets a stop codon (triplets UAA, UAG or UGA). Then, eukaryotic release factors (eRFs) are recruited to separate the polypeptide from the tRNA. Finally, recycling of the ribosome starts with 60S eviction and deacetylated tRNA and mRNA dissociation from the 40S subunit (Dever and Green, 2012) (Fig. I11).

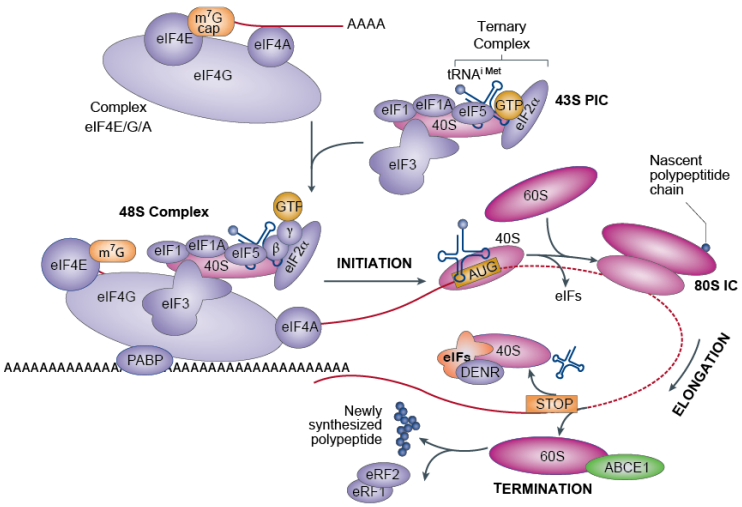


Figure I11: Overview of the translation process. Schematic representation of the three phases of translation (initiation, elongation and termination) focused on PIC assembly during initiation and ribosome disassembly during termination. 43S PIC is formed by the small ribosomal subunit, eukaryotic initiation factors (eIFs 3, 1, 1A and 5) and the ternary complex (Met-tRNA$_i$ [aminoacylated tRNA loaded with methionine], eIF2 and guanine triphosphate (GTP). Interaction between eIF4G and eIF3 allows the interaction between the PIC and the mRNA forming the 48S complex. Ribosome recycling starts with 60S eviction mediated by the ABCE1 (ATP-binding cassette subfamily E member 1) while the deacetylated tRNA and mRNA are dissociated from the 40S subunit by DENR (density-regulated protein) and further eIFs. Adapted from Bhat et al., 2015.

3.2. Transcriptional regulation of ribosomal protein genes

The coding sequences of the RP genes (RPGs) are highly conserved along evolution; however, the promoter features and the machinery involved in their transcriptional regulation is not conserved. In prokaryotes, RPGs form operons, in which the expression of multiple genes is controlled by a single promoter. Besides, some RPs can negatively regulate their own translation (McGary and Nudler, 2013; Nomura et al., 1984). In the case of eukaryotes, the situation is much more complex with multiple genes widely scattered along the genome (Hu and Li, 2007; Yoshihama et al., 2002).

In *S. cerevisiae*, the main elements involved in RPG transcriptional regulation have been characterized (Bosio et al., 2017). In addition, several regulatory mechanisms connecting RPG expression to signaling cascades have been identified, mostly the Target of Rapamycin complex 1 (TORC1) pathway, which is in charge of sensing nutrients and mitotic factors. RPG promoters are classified in two categories depending on the presence of the High mobility group protein 1. In both cases, Repressor activator protein 1 (Rap1) acts as a pioneer element required for the

binding of other TFs. Rap1 and Forkhead-like 1 (Fhl1) are present constitutively at the RPG promoter and the co-activator Interacts with Forkhead 1 (Ifhl1) is recruited to form a complex with Fhl1 only upon activating signals (Knight et al., 2014) (Fig. I12). A small subset of RPGs (6%) appear to be Rap1-independent and to require instead ARS-binding factor 1, whose association with target promoters is dynamic and responsive to environmental stimuli (Fermi et al., 2016).

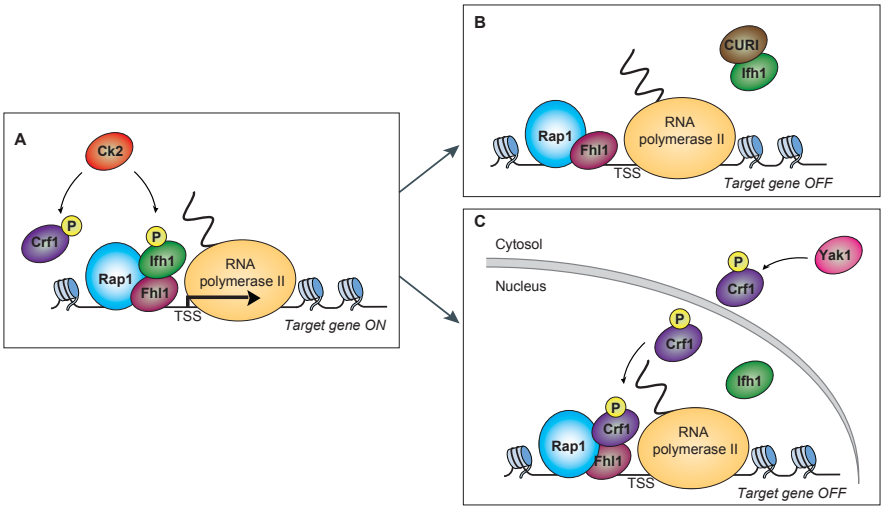


Figure I12: Mechanism of transcriptional regulation of RPGs in *S. cerevisiae*. (A) Transcription activation occurs when Ifh1 is recruited to the promoter. Ck2 favors transcription by phosphorylating Ifh1 and the repressor Crf1. **(B)** CURI complex sequesters Ifh1 leading to gene silencing. **(C)** Yak1 phosphorylates Crf1, allowing its nuclear translocation and competition for Fhl1 binding sites. For more details, see full text.

Two main regulatory mechanisms have been described (Fig. I12). One of them relies on the CURI complex, composed by Ck2 and the rRNA processing proteins U3 snoRNA associated protein 22 and rRNA protein 7, and it has been suggested to be a mediator between rRNA synthesis and RP production since the CURI complex sequesters Ifh1 to repress RPGs when rRNA transcription is impaired (Albert et al., 2016). The other one depends on TORC1 inactivation, and Co-repressor with Fhl1 (Crf1) competes with Ifh1 for the binding to Fhl1 upon nuclear translocation induced by Yak1 phosphorylation (DYRK homolog in *S. cerevisiae*) (Martin et al., 2004). In addition, Ck2 phosphorylates both Ifh1 (Thr681) and Crf1 (Thr348) in their Fhl1 binding domains. Ifh1 phosphorylation facilitates the Fhl1-Ifh1 interaction whereas phosphorylation of Crf1 inhibits the repressive role of this factor (Kim and Hahn, 2016).

Despite this information, a clear molecular mechanism for the regulation of the RPGs remains unknown in mammals. Only some analogies at the

functional level have been extracted: for instance, Split finger protein 1 has been suggested to be the functional analog of c-Myc in *S. cerevisiae* since it is targeted by the TORC1 complex, and it links directly this signaling pathway with the transcription of canonical and auxiliary RPGs (Lempiainen and Shore, 2009).

Given the lack of evolutional conservation of the yeast factors, studies have been focused on searching TF binding sites within the promoters of the RPGs in higher eukaryotes. The studies have revealed that there is not a unique motif common to all metazoans RPGs; by contrast, distinct motifs are found differentially enriched within the RPG promoters, suggesting that they are regulated by different combinations of TFs in different organisms and physiological situations (Perina et al., 2011). The DNA sequences considered as the ones most enriched in mammalian RPG promoters are described below:

- ***Tract of polypyrimidine / TCT motif***: the motif appears in RPG promoters and other genes whose products are related to translation. It is found close to the TSS and it is thought to play a dual role both in transcription and translation initiation (Perina et al., 2011). The TCT-motif is recognized preferentially by TBP-related factor 2 (TRF2) in *Drosophila* (Baumann and Gilmour, 2017; Wang et al., 2014). TRF2 has been proposed to regulate RNAPII transcription by targeting TAF1 to subsets of TATA-less promoters in a TFIIB-independent manner (Baumann and Gilmour, 2017; Zehavi et al., 2015). The human homolog is named TBP-like 1, but whether there is functional conservation has not been investigated yet.
- ***TATA-like sequences:*** the RPG promoters are generally classified as "TATA-less", but they do contain A/T-rich sequences and TBP may be present at human RPG promoters (Perina et al., 2011; Perry, 2005).
- ***GABP motif:*** the motif is found upstream the TSS in 66% of human and mouse RPG promoters (Perina et al., 2011; Perry, 2005). GA-binding protein (GABP) belongs to the ETS family, and it forms a heterotetramer to carry out its function as a transcriptional activator of genes required for the regulation of cell cycle, differentiation, protein synthesis and cellular metabolism (Rosmarin et al., 2004).
- ***YY1 motif:*** the motif is predominantly found downstream the TSS of RPGs (Perina et al., 2011; Perry, 2005). Yin Yang 1 (YY1) is a ZF-containing protein, which is involved in multiple mechanistic roles as an activator, a repressor or a chromatin modifier (Khachigian, 2018). Interestingly, a *de novo* motif named M1 has been found as the most highly enriched binding site in mouse and human RPG promoters (Roepcke et al., 2006). The motif partially overlaps with that of YY1 but

the authors claim that it may have additional roles in transcription initiation and, possibly splicing, based on its constant location in the first exon of the RPGs.

- ***Sp1 motif:*** the motif is ubiquitously found in genes implicated in processes such as cell growth, differentiation, apoptosis, development and carcinogenesis (Beishline and Azizkhan-Clifford, 2015). The associated protein Specificity protein 1 belongs to the family of Sp/Kruppel-like TFs and is considered to be an activator of genes containing putative CG-rich sites in their promoters (Safe et al., 2014).
- ***Myc motif:*** the motif is commonly found in genes involved in a wide range of cellular processes (cell proliferation, cell growth, apoptosis, self-renewal) and therefore c-Myc protein targets a huge plethora of genes (Kress et al., 2015), including genes encoding direct components of the ribosome as well as factors required for rRNA processing (van Riggelen et al., 2010).
- ***M4-motif:*** the motif was detected as enriched in mouse and human RPG promoters (Roepcke et al., 2006; Wyrwicz et al., 2007). The sequence was recently acknowledged as a *de novo* KAISO motif (Raghav et al., 2012), and it corresponds to the DYRK1A-motif (Di Vona et al., 2015). Yamashita and colleagues considered the M4-motif a DNA replication element (DRE)-like sequence in the human genome, potentially able to be recognized by human DRE factor/zinc finger BED-type containing 1 (hDREF/ZBED1, from herein hDREF) (Yamashita et al., 2007). In fact, the DRE motif was found to be highly enriched in RPG promoters in flies (Ma et al., 2009), and DREF has been proposed to target TRF2 to a subset of them (Baumann and Gilmour, 2017).

3.3. Biological impact of a differential regulation of ribosomal proteins

3.3.1. Specialized ribosomes

Ribosome structure and function have been known for decades. Although it is composed of many different factors, it has been considered as an automatic machine, acting in a constitutive manner. In the last few years, however, it has been observed that ribosomes present differential composition of their elements. This fact it is thought to play a role in specialized translation of mRNAs, thus adding another layer of regulation to gene expression (Shi and Barna, 2015; Xue and Barna, 2012). There are different sources of variability leading to distinct ribosomes: i) post-transcriptional modifications of rRNA, ii) post-translational modification of RPs, and iii) differential composition either in auxiliary or in core RPs (Shi and Barna, 2015). Regarding the last point, there is evidence pointing to

the existence of differential composition of RPs within the ribosomes, which might be the result of the regulation of RPG transcription.

Firstly, in 80-90% of the RPs the molar ratio of their transcripts varies around three-fold. This variation implies that there is unequal regulation of the expression of RPGs. In fact, most RPGs show differential expression in mouse embryonic tissues (Kondrashov et al., 2011). In addition, mass spectrometry (MS) studies have shown that the stoichiometry of RPs forming the ribosome changes depending on the number of ribosomes per mRNA (Slavov et al., 2015). Distinct regulatory mechanisms have been identified, such in the case of RPL41/eL41, which is much more transcribed than other genes but very inefficiently translated, or RPS24/eS24, which contains an intron alternatively spliced during the differentiation of embryonic stem cells (ESCs) into myoblasts (Gupta and Warner, 2014).

Secondly, in yeast, plants and flies, it is common to find RPs encoded by more than one copy, so their resulting proteins might play subtly different roles despite having high sequence homology (Xue and Barna, 2012). Mammalian RPs are mostly encoded by single copy genes. However, there are cases in which the paralogs of some RPGs are expressed fulfilling important roles under specific physiological situations. For instance, the mRNAs encoding for RPL10/uL16, RPL39/eL39 and RPL3/uL3 and their corresponding paralogs (RPL10L, RPL39L and RPL3L) are alternatively transcribed in a tissue specific manner (Wong et al., 2014). An example of paralog-specific functions is the case of the pair RPL22/eL22 and RPL22L1, which perform antagonistic roles in the context of the Mothers against decapentaplegic homolog 1 pathway during the hematopoietic development in *Zebrafish* (Zhang et al., 2013).

Based on experimental evidence, it has been suggested that the function of specialized ribosomes is to translate specific pools of mRNAs. For instance, RPL38/eL38 is specifically needed for the translation of homeotic *Hox* genes in mice embryos (Kondrashov et al., 2011), by allowing the recognition of an internal ribosome entry site (IRES) element required for the expression of these genes (Xue et al., 2015). As a consequence, the absence of RPL38/eL38 affects the translation of *Hox* genes, leading to axial skeleton patterning alterations. Likewise, RPL10A/uL1 and RPS25/eS25 direct different sets of ribosomes to the translation of specific mRNAs (Shi et al., 2017).

3.3.2. p53 cell cycle arrest and extra-ribosomal functions

Another layer of complexity comes from the fact that some RPs have extra-ribosomal functions. The most known one is related to p53-

dependent cell cycle arrest. When RPs such as RPL5/uL18 or RPL11/uL5 are not incorporated correctly into ribosomes, they bind murine/human double minute 2 (MDM2/HDM2) and inhibit p53 degradation (Bursac et al., 2012). On the other hand, RPs may have important roles as RNA-binding proteins in rRNA processing and in the translational control of their transcripts. That would be the case of RPS13/uS15, which inhibits the splicing of its own transcript (Malygin et al., 2007), or RPL13/uL13, which regulates the translation of specific mRNAs as part of a non-ribosomal complex (Mazumder et al., 2003).

3.4. Ribosomes and disease

The imbalance in the production of RPs drives the development of aberrant phenotypes. Since protein synthesis is directly linked to cell growth, the most common phenotype due to loss of function mutations in RPGs is a reduction in body size (Fig. I13A). This phenotype was first described as *minute* in *D. melanogaster* and it comprises more than 50 mutations scattered throughout the genome, many of which are located in RPGs (Shi and Barna, 2015). The fact that reduced *RPS3/uS3* mRNA levels correlates with the severity of the phenotype suggests that the sensitivity for reduced levels of particular RPG products may be different (Saeboe-Larssen et al., 1998). Another common outcome found in many mutated RPs is bone marrow defects (Fig. I13A). The most studied one is the Diamond-Blackfan anemia, commonly associated to *RPS19/eS19* haploinsufficiency, although it has also been linked to mutations found in at least ten more RPGs (Xue and Barna, 2012).

A

	Zebrafish	Mouse	Human
Anemia	*rpS19 (es19), rpl11 (el11), rps29 (us14)*	*Rps6 (es6), Rps19 (es19)*	Diamond-Blackfan anemia: *RPS19 (eS19), RPL5 (uL18), RPL11 (uL5), RPS10 (eS10), RPL35A (eL33), RPS26 (eS26), RPS24 (eS24), RPS17 (eS17), RPS7 (eS7), RPL26 (uL24), RPS29 (uS14)*
Small body size	*rps15a (us8)* and 12 other RP gene mutations	*Rps7 (es7), Rps19 (es19), Rpl24 (el24), Rpl27a (ul15) Rpl29 (el29), Rpl38 (el38), Rplp1*	Short stature: *RPS19, RPuL18, RPuL5*, etc.

B

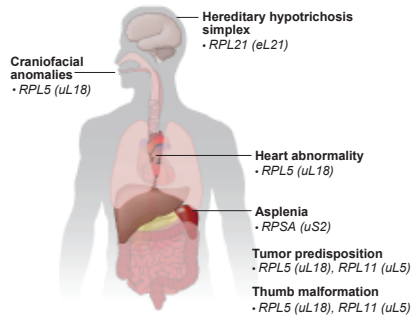


Figure I13: Ribosomal protein loss-of-function mutations in human. Brief description of the phenotypes and their corresponding RPG mutations are shown. **(A)** Most commonly shared phenotypes conserved in zebrafish, mouse and human. **(B)** Highly specific diseases indicating the organ that displays the phenotype. Adapted from Shi and Barna, 2015.

Defects in ribosome functionality are also connected to cancer. Some translation factors (eIF4E/G, eIF3 complex or eEF1α2) as well as TFs (i.e. c-Myc) and RNAPIII components boosting ribosome biogenesis are frequently overexpressed in tumors; unexpectedly, loss-of-function (LoF) mutations in RPGs have been also found in a certain tumor types (Bustelo and Dosil, 2018). These results suggest that increased translational rates might favor tumorigenesis, but also that the alterations in ribosome biogenesis might contribute to shape the tumor proteome.

Finally, it is worth noting that there are RPGs LoF mutations leading to very specific phenotypes. For example, mutations in the *RPSA/uS2* cause lack of spleen or mutations in *RPL21/eL21* are responsible for hereditary hypotrichosis simplex (Shi and Barna, 2015) (Fig. I13B). All these observations reinforce the hypothesis of a high specialization of ribosomes.

Objectives

The DYRK1A protein kinase has been shown to be recruited to specific genomic regions. An interesting subset of DYRK1A genomic *loci* are related to cell growth, including the promoters of RPGs. Notably, some cell lines (including the cell model used in this thesis) present a reduction in cell size upon DYRK1A silencing, possibly due to a loss in cell mass (Di Vona, 2013). It is therefore possible that DYRK1A exerts a role in translation by regulating the transcript levels of RPGs. Very little is known on how RPG expression is regulated in mammals. Therefore, this research will not only provide novel molecular explanations to the DYRK1A-related phenotypes, but it can also contribute to increase our understanding on the regulation of the expression of RPGs.

The objectives of this Thesis work are the following:

I. To characterize the mechanism of DYRK1A recruitment to RNAPII-associated genomic regions, and in particular, to RPG promoters. Specific questions to be answered are: Are there any DYRK1A binding partner at the chromatin level? If so, how does DYRK1A cross-talk with its partners in this context?

II. To elucidate whether DYRK1A regulates the expression of RPGs and consequently affects protein synthesis.

Materials and Methods

1. Plasmids

All plasmids used in this work were generated by cloning, purchased to companies or kindly provided by other laboratories. Their identities were checked by DNA sequencing.

1.1. Backbone vectors

- pCDNA-3: mammalian expression vector (Invitrogen).
- p3XFLAG-CMV-7: mammalian expression vector to express N-terminally Flag-tagged fusion protein under the cytomegalovirus (CMV) promoter (Sigma).
- pGEX-4T3: expression vector for bacterial cells with an isopropyl β D-1-thiogalactopyranoside (IPTG) inducible N-terminal glutathione S-transferase (GST) tag (Amersham Biosciences).
- pGEX-6P1: expression vector for bacterial cells with an IPTG inducible N-terminal GST tag and a preScission protease cutting site (Amersham Biosciences).
- pGEX-6P3: expression vector for bacterial cells with an IPTG inducible N-terminal GST tag and a preScission protease cutting site (Amersham Biosciences).
- pEGFP-C1: expression vector for green fluorescent protein (GFP; Clontech).

1.2. Bacterial expression plasmids for tag-fusion proteins

- pGST-DYRK1A: expression plasmid encoding human DYRK1A (754 aa isoform; NM_130436) with an N-terminal GST tag (Alvarez et al., 2007).
- pGST-DYRK1A^{KR}: expression plasmid encoding for an N-terminally GST-tagged catalytic inactive version of DYRK1A with a point mutation in the lysine residue 179 within the ATP binding pocket (Alvarez et al., 2007).
- pGST-KAISO: expression plasmid encoding murine KAISO (NM_020256) with an N-terminal GST tag. The plasmid was kindly provided by Dr. Mireia Duñach (Universidad Autónoma de Barcelona, Spain) (del Valle-Perez et al., 2016).
- pGST-E2F1: expression plasmid encoding human transcription factor E2F-1 (NM_005225) with an N-terminal GST tag (Ormondroyd et al., 1995).

The following plasmids were generated by Gibson cloning in collaboration with Krisztina Arató (de la Luna's group, CRG). The DNA fragments corresponding to the aa indicated were polymerase chain reaction (PCR)-

amplified from the pFlag-BRCA1 plasmid (MM1.3), and then ligated into the BamHI sites of pGEX-6P1 vector to produce N-terminally tagged GST fusion proteins (see section MM2.2, for specific information):

- pGST-BRCA1-N: N-terminal fragment of BRCA1 corresponding to aa 1-316.
- pGST-BRCA1-M1: BRCA1 fragment corresponding to aa 310-589.
- pGST-BRCA1-M2: BRCA1 fragment corresponding to aa 582-889.
- pGST-BRCA1-M3: BRCA1 fragment corresponding to aa 886-1205.
- pGST-BRCA1-M4: BRCA1 fragment corresponding to aa 1201-1498.
- pGST-BRCA1-C: C-terminal fragment of BRCA1 corresponding to aa 1490-1863.

1.3. Mammalian expression plasmids for tag-fusion proteins

- pFlag-BRCA1: expression plasmid encoding human BRCA1 (1863 aa isoform; NP_009225) with an N-terminal Flag-tag. It was kindly provided by Dr. Richard Baer (Columbia University Medical Center, New York, US) (Choudhury et al., 2004).

1.4. Plasmids for lentivectors preparation

- pCMV-VSV-G: lentiviral packaging vector that expresses the vesicular stomatitis virus G envelope protein (Stewart et al., 2003). Obtained from Addgene (Plasmid #8454).
- pCMV-dR8.91: second generation packaging plasmids containing gag, pol, and rev genes proceding from human immunodeficiency 1 virus (Zufferey et al., 1997), kindly provided by Dr. Trono (Laboratory of Virology and Genetics, École Polytechnique Fédérale de Lausanne, Switzerland).
- pshControl: MISSION® pLKO.1-puro Non-Target shRNA Control plasmid (Sigma #SHC016).
- pshDYRK1A-1: MISSION® pLKO.1-puro shRNA DYRK1A plasmid (Sigma #TRCN0000022999).
- pshDYRK1A-2: MISSION® pLKO.1-puro shRNA DYRK1A plasmid (Sigma #TRCN0000199464).
- pshCHD2-1: MISSION® pLKO.1-puro shRNA CHD2 plasmid (Sigma #TRCN0000021335).
- pshCHD2-2: MISSION® pLKO.1-puro shRNA CHD2 plasmid (Sigma #TRCN0000021334).
- pshKAISO: MISSION® pLKO.1-puro shRNA KAISO plasmid (Sigma #TRCN0000017838).
- pshBRCA1: MISSION® plKO.1-puro shRNA BRCA1 plasmid (Sigma #TRCN0000039833).

2. Techniques for DNA manipulation

2.1. DNA purification and sequencing

For small-scale purification, DNA was extracted from bacterial minicultures (3 ml) using the QIAGEN Plasmid Mini Kit (Qiagen) following manufacturer's instructions. Plasmid DNA purification for cell transfections required higher volume bacterial cultures (200 ml) and the QIAGEN Maxi Plasmid Kit (Qiagen) was used according to manufacturer's instructions.

DNA sequencing followed the Sanger method (Sanger et al., 1977), and it was carried out at the Genomic Sequencing Service (Universitat Pompeu Fabra-PRBB, Barcelona). Briefly, around 150-300 ng of DNA was incubated with 0.25 μl of Big-Dye reagent, 2 μl of 5x sequencing buffer (Ready Reaction Cycle Sequencing Kit; Applied Biosystems), and 3.2 pmol of primer (final vol, 10 μl). The DNA amplification was performed under the following conditions: denaturation was at 94ºC for 3 min, followed by 39 cycles of 10 s/95ºC, 30 s/55ºC, and 4 min/60ºC. In the Genomic Sequencing Service, the samples were purified using Sephadex G-50 DNA Grade (GE Healthcare Life Sciences) and analyzed with the 96-capillary Applied Biosystems 3730xl Genetic Analyzer. 4Peaks free software (Nucleobytes) was used to view and analyze the sequences.

2.2. Gibson cloning

Gibson cloning procedure (Gibson et al., 2009) was performed in collaboration with K. Arató to generate 6 fragments derived from human BRCA1 coding sequence fused to a GST tag in their N-terminal end. DNA fragments were PCR-amplified from pFlag-BRCA1 (MM1.3). The primers contain *BamHI* cutting sites that were incorporated into the flanks of the PCR products (Table MM1).

The PCR reaction was performed with 5 ng of template, 0.4 μM of forward and reverse primers, 200 μM of deoxyribonucleotides (dNTPs), 2.5 μl of 10x reaction buffer and 1 μl of Taq polymerase (5 U/μl, Expand High Fidelity System, Roche) in a final volume of 25 μl. The conditions of the PCR program were: initial denaturation at 94ºC for 2 min, followed by 30 cycles of 15 s/94ºC, 30 s/55ºC and 2 min/72ºC), with and additional final extension step (10 min, 72ºC).

The donor vector pGEX-6P1 was digested with the *BamHI-HF* restriction enzyme (New England Biolabs) according to manufacturer's conditions and it was fractionated in a 1% (w/v) agarose gel (Ecogen) prepared in 1x Tris/Borate/EDTA (TBE) buffer. The linear plasmid was purified from the melted agarose with the QIAQuick Gel Extraction Kit (Qiagen) following

manufacturer's indications. The Gibson reaction was performed with 150 ng of the purified vector, 1 μl of the PCR product and 10 μl of 2x Gibson reaction mixture (CRG Biomolecular Screening and Protein Technologies Unit) in a final volume of 20 μl for 1 h at 50ºC.

Table MM1: Sequences of oligonucleotides used for PCR amplification of Gibson cloning products.

Fragment[a]	Forward primer	Reverse primer
N: 1-316	GGGGCCCCTGGGATCCATG GATTTATCTGCTCTTCG	GGAATTCCGGGGATCCTCAG CTCCTTGCTAAGCCAG
M1: 310-589	GGGGCCCCTGGGATCCCAG CCTGGCTTAGCAAGGAG	GGAATTCCGGGGATCCTCAG CTGCTTATAGGTTCAG
M2: 582-889	GGGGCCCCTGGGATCCACG AAAGCTGAACCTATAAG	GGAATTCCGGGGATCCTCAA GAGTGGGCAGAGAATG
M3: 886-1205	GGGGCCCCTGGGATCCTCT GCCCACTCTGGGTCC	GGAATTCCGGGGATCCTCAC CCTCTTCGGTAACCC
M4: 1201-1498	GGGGCCCCTGGGATCCGGT TACCGAAGAGGGGCC	GGAATTCCGGGGATCCTCAA GGGGATGACCTTTCC
C: 1490-1863	GGGGCCCCTGGGATCCGAA CCAGGAGTGGAAAGGTC	GGAATTCCGGGGATCCTCAG TAGTGGCTGTGGGGG

[a] The aa numbers correspond to human BRCA1 with Acc. No. NP_009225.

For bacterial transformation, 1 μl of the final product was added to 45 μl of competent *Escherichia coli* XL-10 Gold bacteria (Stratagene) and the heat shock method was used (0ºC/15 min, 42ºC/90 s, 0ºC/5 min). Cells were allowed to grow in 200 μl of SOC medium (20 g/l tryptone, 5 g/l yeast extract, 10 mM NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) for 45 min at 37ºC with 220 rpm agitation. The cultures were plated on Lysogeny Broth (LB) plates (10 g/l tryptone, 5 g/l NaCl, 5 g/l yeast extract, 12 g/l agar) supplemented with ampicillin (100 μg/ml) at 37ºC overnight.

Finally, some colonies were picked in 100 μl of LB media and plasmid validation was checked by colony-PCR: 2 μl of inoculated LB, 0.2 μM of forward and reverse primers, 200 μM of dNTPs, 2 μl of 10x reaction buffer and 0.1 μl of Taq polymerase in a final volume of 20 μl. The conditions of the PCR program were: initial denaturation step (1 min, 94ºC); 30 cycles of denaturation step (30 s, 94ºC), primer annealing (30 s, 55ºC) and extension step (2 min, 72ºC); and additional final extension step (7 min, 72ºC). The PCR products were fractionated in a 1% agarose gel, and the positive samples were confirmed by DNA sequencing. The plasmid DNA was transformed into *E. coli* BL21 (DE3) pLysS B F' bacteria (Stratagene) for expression of the fusion proteins.

3. Cell Culture

3.1. Cell lines

The following eukaryotic cell lines used in this work:

- HEK-293T: epithelial cell line derived from human embryonic kidney transformed with the SV40 virus large T antigen.
- T98G: fibroblastic cell line derived from human glioblastoma.
- HeLa: epithelial cell line derived from human cervical adenocarcinoma.
- Kc167: spontaneously immortalized cell line derived from disaggregated *D. melanogaster* embryos.

The HEK-293T, T98G and HeLa cell lines were obtained from the American Type Culture Collection (www.atcc.org); Kc167 cells were kindly provided from Guillaume Fillion's laboratory (CRG). Human cells grew in Dulbecco's Modified Eagle's Medium (DMEM, GIBCO, Thermo Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS, GIBCO, Thermo Fisher Scientific) and antibiotics (100 U/ml penicillin and 100 μg/ml streptomycin, Invitrogen) at 37ºC and in a 5% CO_2 atmosphere. Fly cells were cultured in Schneider's *Drosophila* medium (Thermo Fisher Scientific) supplemented with 5% FBS (Hyclone) at 25ºC.

For depletion of protein expression, small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs) were used. They were delivered by lipofectamine transfection (MM3.2) and lentiviral infection (MM3.3), respectively. Experiments only proceeded when the knock-down efficiency was higher than 50%. The expression levels were assessed by Western blot (MM4.2).

3.2. Cell transfection

Cells were transfected with or 8-16 μg of DNA (100-mm plates), depending on the production/stability of the expressed protein, by the calcium phosphate precipitation method (Graham and van der Eb, 1973). The optimal pH of the HEPES buffer for each cell line was assessed by transfection with a GFP-expressing construct and quantification of fluorescent cells. The DNA-calcium phosphate precipitate was removed after 16 h by washing the cells with Phosphate-buffered saline (PBS) and adding fresh complete medium. Cells were processed at 48 h after transfection according to the purpose of the experiment.

For siRNAs delivery, Lipofectamine® 3000 transfection kit (Thermo Fisher Scientific) was used according to manufacturer's instructions. Briefly, 30 nM siRNA was added to T98G cells incubated in DMEM without antibiotics. The cells were washed with PBS after 16 h and fresh complete

DMEM was added. After 24 h, the cells were processed according to the purpose of the experiment. siRNAs were obtained from GE Healthcare Dharmacon Inc.

- siControl: ON-TARGETplus Non-Targeting Pool (D-001810-10) with the target sequences UGGUUUACAUGUCGACUAA, UGGUUUACA UGUUGUGUGA, UGGUUUACAUGUUUUCUGA, and UGGUUUAC AUGUUUUCCUA.
- siKAISO: ON-TARGETplus human ZBTB33 (10009) SMARTpool (L-019982-00) with the target sequences GAUGGAAGGGUCUAUUA UA, GUACAGCGCAGGAUGGUAA, UAAAGUCAGUUCA UAGUCA and UAAAUCAGGGAGCUUAUUA.

3.3. Preparation of lentiviral stocks and infection

The CRG has all the permits to work with lentivirus in a biological contention level 2 (A/ES/05/I-13 and A/ES/05/14), and the activity of the group has been positively evaluated by the CRG Biological Safety Committee. To generate a viral stock, HEK-293T cells were seeded at a density of $2.5x10^6$ in 100-mm plates and transfected with 3.4 μg of pCMV-VSV-G envelope plasmid, 6.3 μg of pCMV-dR8.91 packaging construct and 9.8 μg of the pLKO-based plasmids by the calcium phosphate precipitation method. Fresh DMEM medium was added 24 h after transfection, and the lentivirus containing supernatant was harvested at 48 h and 72 h after transfection. The two supernatants were pooled together, spun at 1,000x*g* for 10 min and filtered through a 0.45 μm filter (Millipore). Viral particles were either added directly to the cells in suspension or concentrated by centrifugation (20,200 rpm for 2 h at 4ºC using a SW32Ti rotor in a Beckman Coulter centrifuge, followed by resuspension of the viral pellet in PBS, aliquoted and stored at -80º C).

For infection, cells were seeded at a density of $8x10^4$ in 35-mm, $4x10^5$ in 100-mm or $1x10^6$ in 150-mm plates. The virus was added to the medium with the cells in suspension, and replaced with fresh complete DMEM 24 h after infection. In some cases, a second round of infection was carried out. The day after, positive selection was performed by adding 1.25 μg/ml of puromycin and incubated for further 48 h (or as long as non-infected control cells were dead). Before processing, cells were allowed to grow in the absence of puromycin for additional 16-24 h.

3.4. FACs analysis of cell cycle parameters

Around $1x10^5$ T98G cells were trypsinized, washed in PBS and pelleted by centrifugation at 1,000x*g* for 5 min. Cell pellets were fixed by adding dropwise 500 μl of cold 70% (v/v) ethanol while gently vortexing the

sample. Fixed samples were stored at -20ºC for, at least, 16 h. Cells were centrifuged and washed with PBS to remove residual ethanol. DNA staining was performed with 500 μl of 4',6-diamino-2-phenylindole (DAPI) solution (1 μg/ml DAPI [Roche], 0.1% [v/v] Triton X-100 in PBS). Cells were analyzed with a LSF II flow cytometer (Becton Dickinson) using the FACSDiva Software v6.1.2 (Becton Dickinson). The cell cycle profile was determined with the program ModFit v3.2 (Verity Software).

4. Techniques for protein analysis

4.1. Preparation of cell lysates

For total cell lysates preparation, $1x10^6$ cells were resuspended in 200 μl of SDS-lysis buffer (25 mM Tris-Cl pH 7.5, 1% [w/v] sodium dodecil sulfate [SDS], 1 mM ethylendiamine tetracetic acid [EDTA], 10 mM sodium pyrophosphate [Na-PPi], 20 mM β-glycerolphosphate), vortexed for 10 s and heated for 10 min at 98ºC (the process was repeated until the lysate got fluid).

Soluble extracts were prepared by lysing the cells in HEPES-lysis buffer (50 mM 4-[2-hydroxyethyl]-1-piperazineethanesulfonic acid [HEPES] pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% [v/v] Nonidet P-40 [NP-40] [Sigma]), supplemented with a protease inhibitor cocktail (cOmplete Mini; Roche Diagnostic) and phosphatase inhibitors (2 mM sodium orthovanadate, 30 mM Na-PPi, and 25 mM FNa); the cell suspension was incubated for 30 min at 4ºC, followed by centrifugation at 13,200 rpm (Eppendorf 5424 Microcentrifuge) for 30 min at 4ºC to separate the lysate from insoluble cellular components.

Subcellular fractionation to obtain the cytoplasmic and the soluble nuclear compartments was performed using NE-PER (Nuclear and Cytoplasmic Extraction Reagent, Thermo Scientific) according to the manufacturer's protocol. To extract the insoluble nuclear/chromatin compartment, the remaining pellet was resuspended in SDS-lysis buffer.

Protein quantification was done with the BCA Protein Assay Kit (Pierce - Thermo Scientific), following manufacturer's instructions.

4.2. Western blot (WB) analysis

Protein extracts were denatured by adding 6x loading buffer (350 mM Tris-HCl pH 6.8, 30% [v/v] glycerol, 10% SDS, 600 mM dithiothreitol [DTT], 0.012% [w/v] bromophenol blue) and incubated at 98ºC for 10 min. The proteins were resolved on SDS-polyacrylamide gels (SDS-PAGE) of different acrylamide percentage (depending on the molecular weight of the protein to study) at 120 V in 1x running buffer (25 mM Tris-base, 200 mM

glycine, 0.1% SDS). Proteins were transferred onto Hybond-ECL nitrocellulose membranes (Amersham Biosciences) at 400 mA for 1 h at 4ºC in 1x transfer buffer (25 mM Tris-HCl pH 8.3, 200 mM glycine, 20% [v/v] methanol). The presence of proteins was checked by Ponceau S (Sigma) staining.

Table MM2: Properties and working dilution of the primary antibodies used.

Primary antibody	Host	Dilution	Commercial brand
BRCA1[a]	Rabbit	1:1000	Bethyl (A300-000A)
BRCA1[a]	Mouse	1:500	Santa Cruz (D-9; 6954)
BRCA1[a]	Rabbit	1:500	Santa Cruz (D-20; sc-641)
BRCA1[a]	Rabbit	1:500	Santa Cruz (I-20; sc-646)
CHD2[b]	Rabbit	1:1000	Abcam (ab68301)
CHD2[b]	Rabbit	1:1000	Cell Signaling (4170)
Cyclin D1	Mouse	1:1000	Cell Signaling (DCS6; 2926)
DREF	Rabbit	2.5 μg	Abnova (H00009189-D01)
DYRK1A	Mouse	1:1000	Abnova (H00001859-M01)
DYRK1A	Mouse	1:1000	Santa Cruz (RR.7; sc-100376)
eIF2α	Rabbit	1:1000	Cell Signaling (9722)
eIF2α-p(Ser51)	Rabbit	1:1000	Cell Signaling (9721)
FLAG	Mouse	1:5000	Sigma (Clone M2; F1804)
GADPH	Mouse	1:10000	Millipore (MAB374)
GST	Mouse	1:1000	Santa Cruz (B-14; sc-138)
KAISO[c]	Mouse	1:1000	Santa Cruz (6F8; sc-23871)
KAISO[c]	Rabbit	1:1000	Bethyl (A303-558A)
Lamin B1	Goat	1:1000	Santa Cruz (C-20; sc-6216)
α-Tubulin	Mouse	1:10000	Sigma (T6199)
Vinculin	Mouse	1:10000	Sigma (V9131)

[a], see Figure R5 for location of the epitopes; [b], see Figure R3 for location of the epitopes; [c], see Figure R4 for location of the epitopes.

Transferred membranes were blocked with 10% (w/v) non-fat milk (Cell Signaling Technologies) or 5% (w/v) bovine serum albumin (BSA) diluted in TBS-T (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1% (v/v) Tween-20 [Sigma]) for 1 h at room temperature The incubation with the primary antibody was performed in 5% non-fat milk in TBS-T or 3% BSA for 16 h at 4ºC (Table MM2). To remove the non-bound primary antibody, membranes were washed three times for 10 min each in TBS-T, and then the secondary antibody (horseradish peroxidase (HRP)-conjugated, Table MM3), was added diluted in 5% non-fat milk or 3% BSA in TBS-T and the incubation proceeded at room temperature for 1 h. After three 10-min washes in TBS-T, the signal in the membranes was revealed with the Western Lightning® Plus ECL (Perkin Elmer) and exposed in a LAS-3000 image analyzer (Fuji PhotoFilm) with the LAS3000-pro software. Band

intensities were quantified with the ImageQuant™ TL software (GE Healthcare Life Sciences). Relative protein levels were calculated using α-tubulin or vinculin as loading controls.

Table MM3: Properties and working dilution of the secondary (HRP-conjugated) antibodies used.

Secondary antibody	Host	Dilution	Commercial brand
Anti-mouse	Rabbit	1:2000-10000	Dako (P0260)
Anti-rabbit	Goat	1:2000-10000	Dako (P0448)
Anti-goat	Rabbit	1:5000	Abcam (ab6741)

4.3. Immunoprecipitation assay

Immunoprecipitation (IP) assays were performed using protein G-Magnetic beads (Invitrogen) for antibodies raised in mouse or protein A-Magnetic beads (Invitrogen) for antibodies raised in rabbit. The beads were incubated with the specific antibody or control immunoglobulins G (IgGs) (Table MM4) for 1 h at room temperature. Pre-clearing of soluble cell extracts (1 μg) or commercial HeLa nuclear extracts (HNE, CIL Biotech; 500 ng) was carried out by incubating the extracts with beads for 1 h at 4ºC. Then, antibody-bound beads and the extracts were incubated for 2 h at 4ºC. The beads were washed three times with washing buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 2 mM EDTA) containing 1% NP-40 plus one additional wash in washing buffer without detergent. Finally, beads-bound proteins were eluted in 2x loading buffer (100 mM Tris-HCl pH 6.8, 200 mM DTT, 4% SDS, 20% glycerol, 0.2% bromophenol blue) for WB analysis (MM4.2) or were used for *in vitro* kinase assays followed by elution and autoradiography analysis (MM4.4).

Table MM4: Properties and amount of the antibodies used in IP.

Antibody	Host	Amount	Commercial brand
BRCA1	Rabbit	2.5 μg	Bethyl (A300-000A)
BRCA1	Mouse	2.5 μg	Santa Cruz (D-9; sc-6954)
BRCA1	Rabbit	2.5 μg	Santa Cruz (D-20; sc-641)
BRCA1	Rabbit	2.5 μg	Santa Cruz (I-20; sc-646)
CHD2	Rabbit	2.5 μg	Abcam (ab68301)
CHD2	Rabbit	2.5 μg	Cell Signaling (4170)
DREF	Rabbit	2.5 μg	Abnova (H00009189-D01)
DYRK1A	Mouse	2.5 μg	Abnova (H00001859-M01)
DYRK1A	Rabbit	2.5 μg	Abcam (ab69811)
DYRK1A	Rabbit	2.5 μg	Sigma (D1694)
FLAG	Mouse	2 μg	Sigma (Clone M2; F3165)
KAISO	Rabbit	2.5 μg	Bethyl (A303-557A)
Normal IgGs	Mouse	2.5 μg	Santa Cruz (sc-2025)
Normal IgGs	Rabbit	2.5 μg	Santa Cruz (sc-2027/Cell Signaling (2729)

4.4. *In vitro* kinase assay

For the *in vitro* kinase assay (*IVK*), two different sources of enzyme were used: GST-DYRK1A expressed in bacteria and purified through glutathione-affinity chromatography or DYRK1A overexpressed with a baculovirus expression system and purified by nickel-affinity chromatography (CRG Biomolecular Screening and Protein Technologies Unit). The enzyme (20-50 ng) was incubated with 50 μM ATP in the presence of 2.5 μCi [γ^{32}P]-ATP (3000 Ci/mmol, Perkin Elmer) and the specific substrate, either bacterially expressed and purified or immunoprecipitated from cell extracts, for 20 min at 30ºC in 20 μl of kinase buffer (12.5 mM HEPES, pH 7.4, 2.5 mM $MgCl_2$, 2.5 mM $MnCl_2$, 0.25 mM DTT). ^{32}P incorporation was determined by eluting the samples from the beads in 2x loading buffer and resolving them by SDS-PAGE. The gel was stained with Coomassie (0.25 % [w/v] Coomassie brilliant blue [Sigma] in 90 % methanol and 10 % [v/v] acetic acid), dried for 1 h at 80ºC and exposed to a film or a Phosphoimager screen (revealed using the GE Typhoon Trio imager). The signal was normalized to the substrate amounts evaluated by WB or Coomassie staining.

For the identification of DYRK1A-dependent BRCA1 phosphosites, bacterially expressed and purified GST-BRCA1 fragments (500 ng) were incubated with GST-DYRK1A (40 ng) in kinase buffer with 1 mM ATP for 20 min at 30ºC.

4.5. Purification of GST-fusion proteins

For the production of GST-fusion proteins, one fresh colony of *E. coli* BL21 (DE3) pLysS B F' (Stratagene) containing the plasmid of interest was inoculated in 5 ml LB medium supplemented with 100 μg/ml ampicillin. The culture grew O/N at 37ºC shaking at 220 rpm. Then, it was re-inoculated (1:50-1:100 dilution) in 100-200 ml LB medium with 100 μg/ml ampicillin. The bacteria culture grew at 37ºC with 220 rpm agitation until an absorbance at 600 nm (A_{600}) of 0.6-0.8 was reached.

Protein production was induced with 0.1 mM IPTG for 8 h at 20ºC for GST-DYRK1A or 4 h at 37ºC for GST-E2F1, GST-KAISO and GST-BRCA1 fragments. For the purification, bacteria were centrifuged at 6,000*xg* for 15 min at 4ºC and pellets were re-suspended in 10 ml of GST-lysis buffer (10 mM Tris-HCl pH 8, 100 mM NaCl, 0.5% NP-40 and 1 mM EDTA) supplemented with protease inhibitor cocktail (cOmplete Mini, Roche Diagnostic). The lysates were sonicated with a Branson Sonifier-250 (3 pulses of 15 s, 10% amplitude) followed by centrifugation at 10,000*xg* for 15 min at 4ºC. For GST-KAISO, 1% (v/v) sarcosyl (Sigma)

and 3% Triton X-100 were added to the lysates before and after sonication, respectively, to avoid protein aggregation.

For the purification of the fusion proteins, the supernatant was incubated for 1 h at room temperature with 100 μl of glutathione-Sepharose beads (Amersham Biosciences) and washed three times in GST-lysis buffer. When required, fusion proteins were eluted by incubation in 100 μl of elution buffer (10 mM L-glutathione reduced [Sigma], 50 mM Tris-HCl pH 8) three times for 10 min rotation at room temperature to a final vol of 300 μl. Unfused GST was used as control in some experiments, and in this case elution was performed with 300 μl of elution buffer in a single step. To remove the free glutathione, proteins were dialyzed against dialysis buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 2 mM EDTA) O/N at 4ºC using dialysis tubing (MWCO 12-14,000, Spectra/Por).

Finally, protein concentration and quality were assessed by resolving the samples by SDS-PAGE and gel staining with Coomassie, using known amounts of BSA as standard.

4.6. Electrophoretic Mobility-Shift Assay (EMSA)

To generate the double strand DNA probes, 50 pmol of oligonucleotide pairs (Table MM5) were annealed by incubation at 98ºC for 5 min in 250 mM Tris-HCl pH 7.4 and then allow the mix to cool down until 25ºC. Radioactive labeling was carried out with 3.5 pmol of either the double or the single stranded probes, 20 μCi [γ^{32}P]-ATP (6000 Ci/mmol, Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs). Unincorporated nucleotides were removed by centrifugation through Sepharose G-25 minicolumns (Pharmacia Biotech AB) leaving the probes at a final concentration of 0.03 pmol/μl in 1x Tris-EDTA (TE) pH 7.5.

For the binding of the probes, 0.5–2 μg of purified GST-fusion proteins were pre-incubated in binding buffer (250 mM Tris-HCl pH 7.5, 250 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 1 mM EDTA, 50% glycerol) for 20 min at room temperature. Then, 0.05 pmol of the labeled probe were added and the incubation proceeded for 10 min at room temperature. For the “competition” assays, a molar excess (4, 10 or 20-fold) of cold wild type or mutated oligonucleotides were added at the same time as the labeled probe and incubation performed for 20 min at room temperature.

The binding reactions were analyzed on 6% native acrylamide gels that were previously pre-run in 0.5x TBE buffer for 30 min at 200 V at 4ºC. Actual run was performed at 150 V for 40 min. Finally, the dried gel was exposed to a film or a Phosphoimager screen revealed using the GE Typhoon Trio imager.

Table MM5: Sequences of oligonucleotides used for EMSA experiments.

Probe name	Forward probe[a]	Reverse probe[a]
DYRK1A-WT	GGCCTAAGAC**TCTCGCGAGA** CACCGTCTAG	CTAGACGGTG**TCTCGCGAGA** GTCTTAGGCC
DYRK1A-MUT	GGCCTAAGAC**AGGTGTACAA** CACCGTCTAG	CTAGACGGTG**TTGTACACCT** GTCTTAGGCC
E2F1-WT	AGGATTTGGCGCGTAAAAGT	ACTTTTACGCGCCAAATCCT

[a]DYRK1A consensus and mutated nucleotides are highlighted in bold.

4.7. ^{35}S-Methionine incorporation assay

T98G cells were seeded in 35-mm plats and incubated for 90 min in free-methionine (Met) DMEM medium (GIBCO, Thermo Fisher) plus 10% dialyzed FBS (GIBCO, Thermo Fisher). Then, 50 μCi [^{35}S]-Met (1,175 Ci/mmol, Perkin Elmer) was added and further incubated for 20 min. Cells were lysed in SDS-lysis buffer and protein extracts were resolved by SDS-PAGE. Equal protein loading was evaluated by WB analysis. The incorporation of [^{35}S]-Met was detected by autoradiography of the dried gel by using a film or a Phosphoimager screen revealed using the GE Typhoon Trio imager.

4.8. Polysome profiling

Polysome profiles were obtained from approximately 1x10^7 T98G cells. Protein synthesis was arrested by incubation with 100 μg/ml cycloheximide (CHX; Sigma) for 5 min at 37ºC. The cells were washed in PBS containing 100 μg/ml CHX, and collected in 1 ml of polysome lysis buffer (10 mM Tris-HCl pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, 1% Triton X-100, 20 mM DTT, 100 μg/ml CHX, 0.25% [v/v] sodium deoxycholate) followed by quick freezing in liquid nitrogen. Finally, cell debris and nuclei were eliminated by spinning the samples at 12,000*xg* for 5 min at 4ºC; the nucleic acid concentration in the supernatants was assessed by A_{260} measurements in a NanoDrop™ (Thermo Fisher).

To fractionate the cell fraction enriched in polysomes, 10-50% linear sucrose gradients (10 ml) were prepared in polysome gradient buffer (20 mM Tris-HCl pH 7.4, 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM DTT, 100 μg/ml CHX) using a Gradient Master 107 (Biocomp) and 14x89 mm polyallomer ultracentrifuge tubes (Beckman). Samples with A_{260}≈10 were loaded onto the top of the gradients, and centrifuged at 35,000 rpm for 3 h at 4ºC in a Beckman SW41Ti rotor. Profiles were taken with continuous monitoring A_{254} using an Econo-UV Monitor (Bio-Rad) and an Econo-Recorder model 1327 (Bio-Rad). Polysome: monosome ratios were obtained by calculating the areas of both fractions using ImageJ software (Schindelin et al., 2012).

4.9. Mass spectrometry analysis

Mass spectrometry (MS) analysis was used to identify DYRK1A-dependent BRCA1 phosphorylation sites (MM4.4). Samples phosphorylated *in vitro* were washed three times with 100 mM ammonium bicarbonate (ABC) buffer, resuspended in 100 μl of 6 M urea-100 mM ABC buffer and incubated for 1 h at 37ºC in the presence of 0.3 mM DTT. Alkylation was carried out with 0.6 mM iodoacetamide (Sigma) for 30 min at room temperature in the dark. Then, proteins were digested in additional 194 μl of 100 mM ABC buffer with 1 μg endoprotease Lys-C (Wako) O/N at 37ºC. Further digestion was performed with 295.8 μl of 100 mM ABC buffer plus 1 μg sequence-grade trypsin (Promega) over day at 37ºC. Finally, digestion was stopped by adding 60.5 μl of formic acid. Peptides were desalted using Ultra Micro Spin Columns C18 (The Nest Group INC) (Rappsilber et al., 2007).

Samples were analyzed using an Orbitrap Fusion Lumos mass spectrometer (MS, Thermo Scientific) coupled to an EasyLC (Thermo Scientific, Proxeon) at the CRG/UPF Proteomics Unit. Peptides were loaded onto the analytical column at a flow rate of 1.5-2 μl/min using a wash-volume of 4 times the injection volume, and were separated by reversed-phase chromatography using a 50-cm column with an inner diameter of 75 μm, packed with 2 μm C18 particles spectrometer (Thermo Scientific). Chromatographic gradients started at 95% buffer A (0.1% [v/v] formic acid in H_2O) and 5% buffer B (0.1% formic acid in acetonitrile) with a flow rate of 300 nl/min and gradually increased to 22% buffer B in 52 min and then to 35% buffer B in 8 min. The column was washed for 10 min with 5% buffer A and 95% buffer B after each sample. The MS was operated in Automated Data Dependent Acquisition (DDA) mode and full MS scans with 1 micro scans at resolution of 120,000 were used over a mass range of m/z 350-1,500 with detection in the Orbitrap. Auto gain control was set to 2e5 and dynamic exclusion to 60 s. In each DDA analysis cycle, following each survey scan Top Speed, ions with charged 2 to 5 above a threshold ion count of 1e4 were selected for fragmentation. Each precursor ion was fragmented both with high-energy collision dissociation (HCD) and electron transfer dissociation (ETD) with collision-induced dissociation supplemental activation of 35%. The quadrupole isolation window was set to 1.6 m/z. Fragment ion spectra produced via HCD were acquired at a normalized collision energy of 28% and both HCD and ETD spectra were acquired in the Orbitrap at 30K. Data were acquired with Xcalibur software (v3.0.63, Thermo Fisher Scientific).

Acquired data were analyzed using the Proteome Discoverer software suite (v2, Thermo Fisher Scientific), and data were searched with Mascot

search engine (v2.4, Matrix Science) against SwissProt *H. sapiens*, where the most common contaminants were added. A precursor ion mass tolerance of 7 ppm at the MS1 level was used, and up to three miscleavages for trypsin were allowed. The fragment ion mass tolerance was set to 20 mmu. Phosphorylation at Ser, Thr and Tyr, oxidation of Met and protein acetylation at the N-terminus were defined as variable modifications. Carbamidomethylation on cysteine (Cys) was set as a fix modification. The phosphoRS node was used for phosphor-site probability determination. The identified peptides were filtered using a false discovery rate (FDR) < 5%.

5. Chromatin immunoprecipitation (ChIP) associated procedures

5.1. ChIP assay

ChIP assays were performed with $8x10^6$ T98G cells per IP. Briefly, formaldehyde was added to the culture medium to a final concentration of 1% (v/v) and incubated for 10 min at room temperature. Crosslinking was quenched with 0.125 M glycine for 5 min. Crosslinked cells were washed twice with cold PBS, resuspended in 1 ml of Lysis buffer I (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% NP-40 plus the protease inhibitor cocktail) and incubated on ice for 10 min. Cells were collected by centrifugation (2,000 rpm for 5 min at room temperature in a Microfuge) and the cell pellet was resuspended in 0.6 ml of Lysis buffer II (1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.0 plus the protease inhibitor cocktail) and incubated for additional 10 min on ice. Chromatin was sonicated to an average size of 0.2–0.5 kb using the Bioruptor (Diagenode). Then, 20 μl of chromatin was reverse-crosslinked at 65ºC O/N and DNA was purified to quantify concentration (using NanoDrop) and to check the size and quality of the DNA fragments (using fractionation in 1% agarose gel).

For each IP, the chromatin amount corresponding to 100 μg of DNA was diluted in 1 ml of IP Buffer (300 mM NaCl, 200 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS and 5% Triton X-100); 10% of the chromatin was used as input-DNA. Samples were incubated with rotation O/N at 4°C with specific antibodies or control IgGs (table MM6). The immunocomplexes were recovered by adding 30 μl of either protein A or protein G sepharose beads (GE Healthcare), depending on the antibody species, and incubated for additional 3 h at 4°C with rotation. Beads were washed with three successive 1 ml-washes of Low salt Buffer (50 mM HEPES pH 7.4, 140 mM NaCl, 1% Triton X-100), one wash with High salt Buffer (50 mM HEPES at pH 7.4, 500 mM NaCl, 1% Triton X-100), one wash with LiCl Buffer (10 mM Tris-HCl pH 8, 250 mM LiCl, 1% NP-40, 1% [v/v]

deoxycholic acid and 1 mM EDTA), and a final wash in TE buffer (all of the washing buffers contained the protease inhibitor cocktail). Elution of DNA from beads was achieved by incubating the samples in Elution Buffer (1% SDS, 100 mM $NaHCO_3$) in two steps of 30 min each at 65ºC with constant agitation (1,000 rpm in a Thermomixer [Eppendorf]); the supernatant was recovered by centrifugation at 3,600 rpm for 5 min at room temperature in a Microfuge. The crosslinking was reverted by an additional incubation of the samples in 200 mM NaCl at 65ºC O/N in agitation at 1,000 rpm. Chromatin-associated proteins were further degraded by adding 1.6 U of protease K (New England Biolabs) for 2 h at 45ºC with 1,000 rpm agitation. The ChIP-DNA was purified by phenol/chloroform extraction followed by ethanol-precipitation. Finally, the DNA was resuspended in 10 μl of nuclease-free water and quantified by using the Qubit® dsDNA HS Assay Kit (Thermo Fisher Scientific) for the ChIP-DNA and the Qubit® dsDNA BR Assay Kit (Thermo Fisher Scientific) for the input-DNA.

Table MM6: Properties and amount of the antibodies used in ChIP.

Antibody	Host	Amount	Commercial brand
Anti-BRCA1	Rabbit	10 μg	Bethyl (A300-000A)
Anti-BRCA1	Mouse	10 μg	Santa Cruz (D-9; sc-6954)
Anti-BRCA1	Rabbit	10 μg	Santa Cruz (D-20; sc-641)
Anti-BRCA1	Rabbit	10 μg	Santa Cruz (I-20; sc-646)
Anti-CHD2	Rabbit	10 μg	Abcam (ab68301)
Anti-CHD2	Rabbit	10 μg	Cell Signaling (4170)
Anti-DREF	Rabbit	10 μg	Abcam (ab48355)
Anti-DYRK1A	Rabbit	10 μg	Abcam (ab69811)
Anti-KAISO	Mouse	10 μg	Santa Cruz (6F8; sc-23871)
Anti-KAISO	Rabbit	10 μg	Bethyl (A303-557A)
Normal IgGs	Mouse	10 μg	Santa Cruz (sc-2025)
Normal IgGs	Rabbit	10 μg	Santa Cruz (sc-2027/Cell Signaling (2729)

5.2. ChIP assay for spike-in normalization

To compare ChIP samples among different conditions more accurately, the use of an external spike-in control was implemented as recommended in (Bonhoure et al., 2014; Orlando et al., 2014). Thus, the differences observed in chromatin occupancy of one factor across different conditions cannot be due to technical issues leading to differential yields but to biological differences.

Briefly, a constant amount of crosslinked chromatin derived from *D. melanogaster* Kc167 cell line was added to each sample in a ratio of 1:500. During the immunoprecipitation step, chromatin was incubated with

the antibody of interest and an antibody recognizing a histone variant specific for *Drosophila* named H2Av (Active Motif, #61686)(Egan et al., 2016). The ChIP procedure was then performed as described in the previous section.

5.3. Quantitative Polymerase Chain Reaction (qPCR)

PCR reactions (10 μl) were performed in triplicate with SYBR Green (Roche) and 10 pmol of specific primers (Table MM7) in 384 well plates using the Roche LC-480 equipment (Roche Applied Science). Denaturation was at 95º for 5 min followed by 50 cycles of 15 s/95º, 20 s/60º, 20 s/72º. The crossing point (Cp) was calculated for each sample using the second derivative maximum method with the Lightcycler 480 1.2 software (Roche), and data analyzed for relative quantification. No PCR products were observed in the absence of template. All primer sets gave narrow single melting point curves, which were checked at the end of each run.

For those cases in which ChIP enrichment is represented as percentage of input, 1/10 dilution of ChIP-DNA and 1/1000 of input-DNA were used for the PCR reactions. For those cases in which ChIP enrichment is represented as fold-change, 0.1 ng of both input-DNA and ChIP-DNA were used for the PCR reactions. When spike-in normalization was used, specific primers amplifying H2Av, *Drosophila* Positive Control Primer Set Pbgs (Active Motif, #71037) and *Drosophila* Negative Control Primer Set 1 (Active Motif, #71028) were used. The enrichment of the positive control was used to calculate a normalization factor to equalize the signal across samples.

Table MM7: Sequences of oligonucleotides used for ChIP-qPCR experiments.

Gene	Forward primer	Reverse primer
ASXL1	AGCATCGCCTCCCAGAAT	CACCGACCTCAGCTAGGAAC
CDK12	TCGCGTTGTTTGATAAGCAG	GATTAAACCGAAACGGCACT
Chr20_COV0[a]	GCTTGGCCAACAAGGTAAAA	CTCTCTGCAACCTCCACCTC
FAM98A	CAAATTTCCGAGTCGTCAGG	AGTGGGACCAAGTTTCTTGC
MRPL16	GCCGGAAGTTGTGTTCACTC	CCTTTGACCGCAGGAACTC
RBM39	AATTTGAGCGGCCGAAGTAT	GAATGGGGGATGGGAATATC
RPS11/uS17	GCTGAAGGCTGGTCACATCT	GGGCACTGTGAAGGACTGAC
RPS15/uS19	TGTGTTTTGCCCTCCAGAC	TCAGAAGAGATCGCTTTGGTC
RPS19/eS19	AACTTTCGCCCTGAGAGAGG	CAGGGGAAAGGGAACGAC
SLC39A13	GAAGTGCGCCCTCTACCC	GTCACAAGGGGCCTGTCC

[a]Oligonucleotides amplifying a random genomic region used as a negative control

5.4. DNA library preparation for ChIP-Seq analysis

Libraries were prepared after checking the quality of the ChIP by evaluating several positive and negative control targets by qPCR (MM5.3). The libraries were generated by using the Ovation® Ultralow Library System V2 (NuGEN) following the manufacturer's instructions. Before sequencing, a quality control was carried out by qPCR analysis of control targets. Libraries were sequenced on an Illumina GAIIx sequencer (HiSeq sequencing v4 chemistry) with a read length of 50 base pair (bp) single end at the CRG Genomics Unit according to established protocols at the Facility. Around 4-6x10^7 reads were obtained for each library with more than 97% aligned reads in all cases.

5.5. Bioinformatic analysis of ChIP-Seq data

The bioinformatic analysis regarding quality control, read mapping and peak identification was done by Roberto Ferrari (Miguel Beato's group at CRG). The analysis was performed basically as described in (Ferrari et al., 2012), with few modifications. Briefly, reads were mapped to the human (hg38) genome assembly using Bowtie software (Langmead et al., 2009), and the Poisson distribution was used to estimate the probability of observing the ChIP counts within a window given the expected counts in the input sample window. All windows with *p*-values less than 1.0x10^{-8} were considered to have significant peaks. The algorithm produced several files that were subsequently used for analysis: BED files contain the coordinates of the significant windows of enrichment and Wiggle (wig) files (chromosome tiling, fixed step) with normalized read counts for the significant windows. When different biological replicates were obtained, the reads coming from them were combined and used for peak calling to improve the coverage of the analysis and dilute possible false positives. Those cases were DYRK1A and KAISO (in untreated cells) and BRCA1 (in silencing experiments).

For peak annotation, an in-house gene and promoter annotation pipeline developed by Sarah Bonnin (CRG Bioinformatics Unit) was used. The script uses data from a DYRK1A consensus position weight matrix (PWM) for the presence of DYRK1A motif (Di Vona et al., 2015), SwitchGear-TSS (www.switchgeargenomics.com) for promoter and TSS definition, and University of California Santa Cruz (UCSC) genome browser software (genome.ucsc.edu) (Kent et al., 2002)_via the UCSC Table Browser (Karolchik et al., 2004) for Ensembl/NCBI gene and transcript annotation. Extra annotation to the final tables was provided using the tool Biomart (Durinck et al., 2005).

The Integrated Genomics Browser v9.0.0 (Nicol et al., 2009) and the UCSC Genome Browser were used for visualization of ChIP-Seq data sets. Further analysis of the ChIP-Seq data was carried out in collaboration with R. Ferrari and Chiara Di Vona (de la Luna's group). Cis-regulatory Elements Annotation System (CEAS) (Ji et al., 2006) was used to create overlaps of significant peaks with genomic annotated regions. Sitepro, as part of the Cistrome Analysis pipeline (cistrome.dfci.harvard.edu/ap), was used to profile average binding (density plot) for defined genomic intervals: plotted values are the -log10 of Poisson *p*-values. Heatmaps were generated using computeMatrix and plotHeatmap tools from deepTools v3.0.0 (Ramirez et al., 2016). Increasing shades of color in the color bar scales stand for higher enrichment and refer to the -log10 of Poisson *p*-values. When indicated, clusters were computed using the *k*-means algorithm. Correlation analysis of selected ChIP datasets was generated using multiBigwigSummary and plotCorrelation tools from deepTools (Ramirez et al., 2016). Correlation coefficients were computed according to the Spearman method v3.0.0 (Spearman, 1987).

For *de novo* detection of motifs associated to the peak regions, the MEME suite v4.12.0 (meme-suite.org) was used (Bailey et al., 2009). The DYRK1A-PWM (Di Vona et al., 2015) was used to define RPG promoters (calculated as -1000 bp/+100 bp from TSS) containing the DYRK1A-consensus using the FIMO package (Grant et al., 2011). The *p*-values for each motif occurrence were converted to *q*-values following the method of Benjamini and Hochberg (Benjamini, 1995).

5.6. Evaluation of alternative ChIP protocols

Some ChIP experiments were performed using the ChIP-IT High Sensitivity Kit (Active Motif) following the manufacturer's instructions. Thus, ChIP data using this system was obtained for DYRK1A, BRCA1 using Bethyl A300-000A antibody and KAISO with Santa Cruz 6F8 antibody. The protocol rendered a gain of DYRK1A, KAISO and BRCA1 associated genomic regions, including those ones co-occupied by the three factors (Fig. MM1).

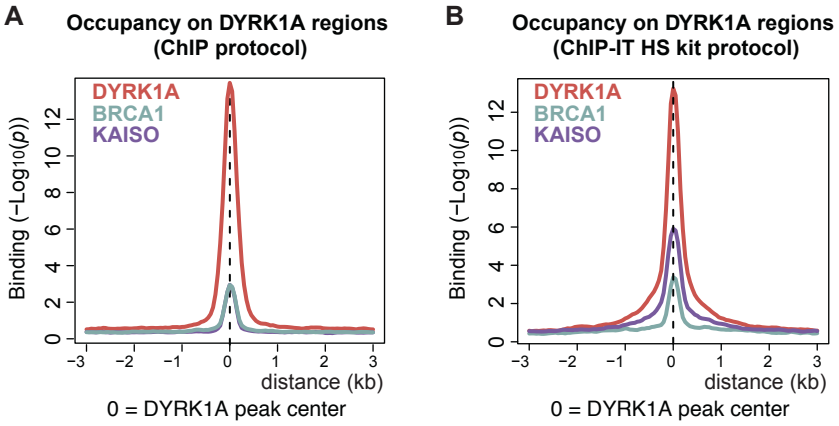


Figure MM1: DYRK1A, KAISO and BRCA1 binding profiles within DYRK1A-associated peaks. Density plots showing the distribution of KAISO and BRCA1 according to the center of the DYRK1A peaks as result of a standard ChIP protocol **(A)** or the ChIP-IT high sensitivity protocol **(B)**. The offset was set to +/- 3 kb from the DYRK1A peak center.

Correlation analysis showed that the output obtained using the kit significantly overlapped with the data obtained with the previous method for the DYRK1A and KAISO profiles (Fig. MM2). By contrast, BRCA1 ChIP-Seq data obtained with the commercial kit correlates less than 50% with the standard protocol (Fig. MM3). Therefore, for DYRK1A and KAISO ChIP-Seqs, the peak calling was carried out combining reads from both datasets, thereby improving the deepness of the analysis. For BRCA1 ChIP-Seq, only the data from the standard method was used, since it gave results that were consistently reproducible across different biological replicates (Fig. MM3).

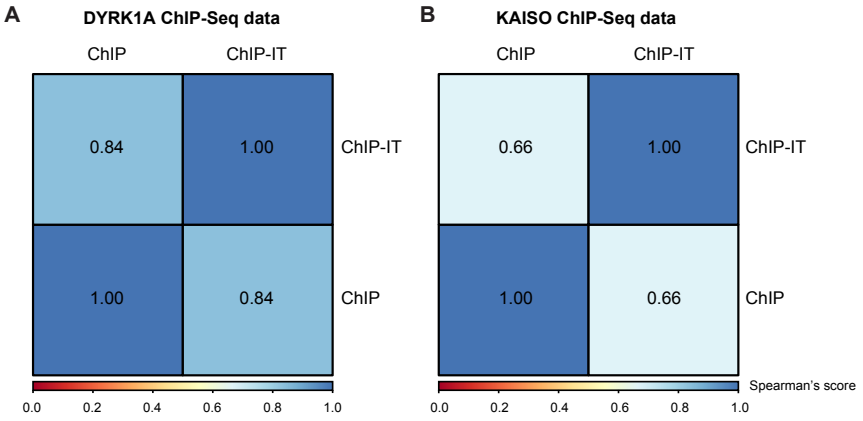


Figure MM2: Correlation analysis of DYRK1A and KAISO ChIP-Seq data obtained with two alternative ChIP methodologies. The graphs show the Spearman's correlation scores for DYRK1A **(A)** and KAISO **(B)** ChIP-Seq data obtained with a standard ChIP protocol (ChIP) or with the ChIP-IT high sensitivity kit protocol (ChIP-IT).

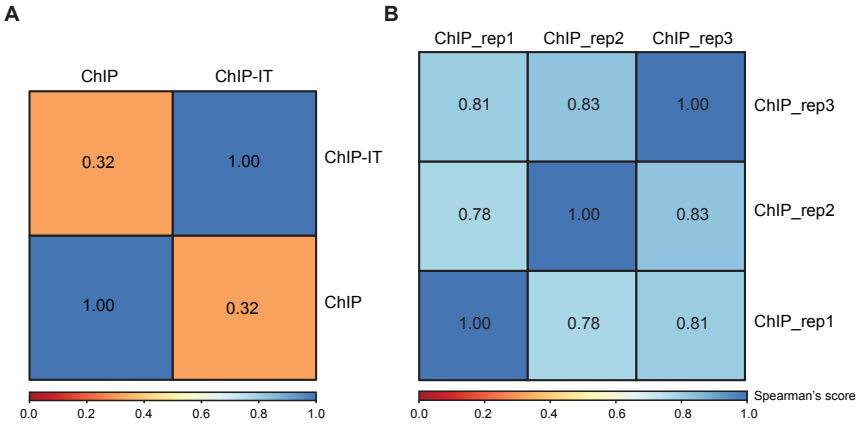


Figure MM3: Correlation analysis of BRCA1 ChIP-Seq data obtained with two alternative ChIP methodologies. (A) Spearman's correlation scores for BRCA1 ChIP obtained with the ChIP high sensitivity kit (ChIP-IT) or with a standard ChIP protocol (ChIP) in untreated cells. **(B)** Spearman's correlation scores obtained for three biological replicates of BRCA1 ChIP in shControl-infected cells, performed following the standardized protocol (ChIP_rep1; ChIP_rep2; ChIP_rep3).

6. RNA analysis

6.1. RNA purification and reverse transcription (RT)

Total RNA was isolated with the RNeasy extraction Kit (Qiagen) or with Trizol (Ambion) following the corresponding manufacturer's instructions. Samples were treated with DNase I (Ambion, 2 U/μl) for 30 min at 37ºC to eliminate the remaining genomic DNA; finally, RNA was quantified with NanoDrop. In some cases, the quality of the RNA was assessed by running samples on a Bioanalizer 2100 (Agilent).

For reverse transcription, 0.5-1 μg RNA was subjected to cDNA synthesis using Superscript II Reverse Transcriptase (RT) (Invitrogen) and random primers as recommended by the manufacturer's instructions.

6.2. RNA preparation for spike-in normalization

Normalization to total RNA might introduce some biases in comparative quantitative analysis, especially in those cases where global transcript up- or down-regulation occurs (Chen et al., 2015; Li et al., 2013; Loven et al., 2012). In this regard, differences in cell volume lead to differential steady-state levels of total RNA to maintain proper RNA concentrations (Padovan-Merhar et al., 2015). Therefore, the traditional normalization method could mask significant biological differences (Fig. MM4).

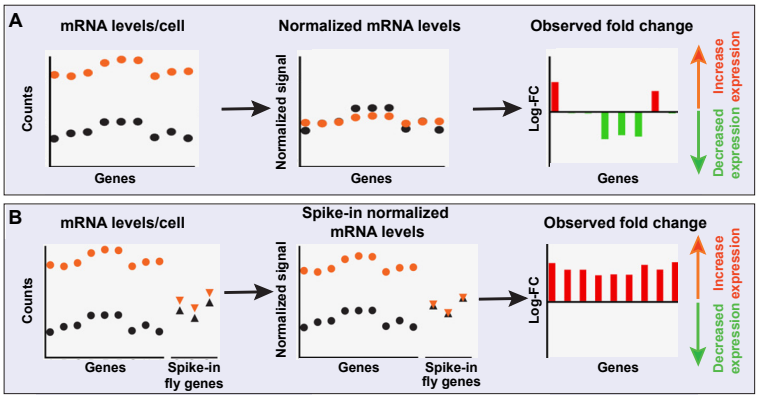


Figure MM4: Alternative normalization methods lead to different interpretation of expression data. Schematic representation of RNA-Seq normalization when the overall levels of mRNA per cell (*y*-axis) are increased in one condition (black) compared to another (orange). The first panel shows the actual relationship of mRNA levels for the two conditions. Increased expression is represented by red bars above the midline, while decreased expression is represented by green bars below the midline. **(A)** Panels depicting the effect of the calculated fold changes based on median normalization. **(B)** Panels depicting the differences when spike-in RNAs are used as standards for normalization (represented by black or orange triangles corresponding to each condition). Adapted from Loven et al., 2012.

Given that DYRK1A-silenced cells are smaller in size compared to control ones (Di Vona, 2013), a spike-in normalization method was implemented for shDYRK1A-infected cells. To this end, equal number of cells for each sample were mixed with a fixed number of cells derived from *D. melanogaster* Kc167 cell line (1:4 ratio), and total RNA was purified as described in the previous section. This allows for the use of *Drosophila* genes as external controls for normalization.

6.3. Quantitative Polymerase Chain Reaction (qPCR)

For RT-qPCR, qPCR was carried out as detailed in MM5.3. Analysis was performed using 1/10 dilution of cDNA as template in 10 μl reaction volumes using SYBR Green, and 10 pmol of specific primers (Table MM8). As negative control, RNA incubated in the absence of RT was included to further monitor contamination of genomic DNA. For spike-in normalization, primers amplifying one cytoskeletal actin gene (*Act42A*) from *D. melanogaster* were used and RNA levels of the samples were normalized according to its expression. Furthermore, the expression of housekeeping genes Eukaryotic elongation factor 1 alpha 1 *(EEF1A1)* and Peptidyl-prolyl-isomerase A *(PPIA)* was assessed as an additional control. In all cases, the values are represented as “fold change” relative to a control sample.

Table MM8: Sequences of oligonucleotides used for RT-qPCR experiments.

Gene	Forward primer	Reverse primer
Act42A[a]	GCGTCGGTCAATTCAATCTT	AAGCTGCAACCTCTTCGTCA
DYRK1A	CCTTGATAGGCAAAGGTTCC	CGCACTTCTATCTGTGCTTG
EEF1A1	AGGTGATTATCCTGAACCATC	GAACGGCGATCAATCTTTTCC
NOL11	TTGGACAAGACGAAAACTCTG	GGTCAGCTGGACGTATTGGT
POLE	ACTTCAACCGCATCATCCTC	ATTCCCAGCATCGAGAGAAA
PPIA	GCCGACCAAAACCGTGTACT	GTCTTTGGGACCTTGTCTGC
RBM28	GATGGAGAGGAACTGGCTCA	TCCCTTCATTCACATCAGAGG
RFC3	ACGTGAACTTTATGGTGTTGG	ACTTGCAATGGTGCTAATTTC
RFC5	GCTTCAGATGACCGAGGAAT	GCGTCTGCTTCATCCAAGAT
RPL14/eL14	GTGCATGCAGCTCACTGATT	TTCAATCTTCTTGGCCCATC
RPL17/uL22	GCTGCACATGCTTAAAAACG	GCGCATCTTAGGTGCTTTGT
RPL21/eL21	AGTTGTTCCTTTGGCCACATA	GGTAACACTTGTGGGGCATT
RPL26/uL24	GGAAAAGGCTAATGGCACAA	TCCTTTCCTACTTGGCGAGA
RPL7A/eL8	AGAAGGCCAAGGGAAAGAAG	AGGTCTCTTTTGGGCTGGAT
RPS15A/uS8	AACCTCACAGGCAGGCTAAA	CGGGATGGAAGCAGATTATT
RPS19/eS19	CAAAGAGCTTGCTCCCTACG	TTCTCTGACGTCCCCCATAG
RPS2/uS5	AAGATCAAGTCCCTGGAGGA	TGCTTCTGCACTGGCATAAT
RPS6/eS6	AGAAGATGATGTCCGCCAGT	CTGCAGGACACGTGGAGTAA
UTP6	TGCATGCTGAAAAACTGAGG	CCAACTCGCCCTTAAGGATT

[a], primers targeting *Act42A* from *D. melanogaster.*

6.4. Library preparation and RNA-Seq

Libraries for RNA-Seq were prepared at the CRG Genomics Unit. Only RNA with RNA integrity number values over 8 was used. Libraries were prepared using the TruSeq Stranded mRNA Sample Prep Kit v2 according to the manufacturer's protocol. Briefly, 1 µg of total RNA was used for poly(A)-mRNA selection using streptavidin-coated magnetic beads; the selected RNA was subsequently fragmented to approximately 300 bp. cDNA was synthesized using Superscript II RT (Invitrogen) and random primers. For the synthesis of the second strand of the cDNA dUTP was used in place of dTTP. dsDNA was subjected to A-tailing and ligation of the barcoded Truseq adapters. Library amplification was performed by PCR using the primer cocktail supplied in the kit. All purification steps were performed using AMPure XP beads (Beckman).

The libraries were analyzed using Agilent DNA 1000 chip to estimate the quantity and check size distribution, and quantified by qPCR using the KAPA Library Quantification Kit (KapaBiosystems) prior to amplification with Illumina's cBot. Libraries were sequenced on an Illumina GAIIx sequencer (HiSeq sequencing v4 chemistry) to a length of 125 bp pair-ended for the case of shDYRK1A experiment and to a length of 50 bp single-ended for the shBRCA1 experiment according to established protocols at the Facility. Around 4-6x10^7 reads were obtained for each

library with more than 97% aligned reads in all cases, and biological triplicates were used for the analysis.

6.5. Bioinformatic analysis of RNA-Seq data

Analysis of RNA-Seq introducing spike-in normalization was performed by Enrique Blanco (Luciano Di Croce's group at the CRG) and C. Di Vona. In brief, the sequences of each RNA-Seq sample were mapped against human and fly genomes (genome assembly versions: hg19 and dm3, respectively) using TopHat (Trapnell et al., 2009) and those reads that could not be uniquely mapped to just one region were discarded. Cufflinks (Trapnell et al., 2010) was run to quantify the expression in FPKMs (fragments per kb of transcript per million mapped reads) of each annotated transcript from the RefSeq catalog of genes on each species (O'Leary et al., 2016). Finally, a local regression method was applied (Loven et al., 2012) to normalize the expression of the human genes according to the fly genes expression to fit the regression line between the two conditions (shControl vs shDYRK1A knocked-down cells).

Mapping of RNA-Seq analysis without spike-in normalization was performed as above without the last step. Differential gene expression was then calculated by normalizing data using the trimmed mean of M-values normalization method (Robinson and Oshlack, 2010) and filtering genes that had >2 average normalized counts per million with the edgeR package (Robinson et al., 2010). Statistical analysis was performed in RStudio by fitting an exact test with the negative binomial distribution for each set of conditions and testing for differential gene expression utilizing the edgeR package (Robinson and Smyth, 2008).

Independently of the normalization method, differential expression was considered for changes with p-value$\leq$0.05, and with fold-changes (FC) above 1.5 and below -1.5 for up- and down-regulated genes, respectively.

7. Databases and other computational tools

Protein and DNA sequences were searched and analyzed by using public databases of the National Centre for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov) and Ensembl (ensembl.org). Bibliography references were queried with the database PubMed from NCBI. Gene expression data for DYRK1A was from GTEx (Genotype-Tissue Expression portal in www.gtexportal.org/home/gene/DYRK1A).

Gene Ontology (GO) enrichment analysis of ChIP-Seq and RNA-Seq data was performed using EnrichR (Chen et al., 2013; Kuleshov et al., 2016) (amp.pharm.mssm.edu/Enrichr). The BioVenn free software

(www.biovenn.nl) was used to overlap gene targets from different datasets (Hulsen et al., 2008).

The Gene Expression Omnibus (GEO) dataset GSE79491 was used to analyze Nur77/NR4A1 ChIP-Seq data obtained from the leukemic Kasumi cell line (Duren et al., 2016). GABP (GSM1010739) and YY1 (GSM1010897) ChIP-Seq data of neuroblastoma SK-N-SH cells as well as CHD2 (GSM935378), KAISO (GSM803392) and BRCA1 (GSM935377) ChIP-Seqs from the lymphoblastoid cell line GM12878 were taken from the ENCODE database (Consortium, 2012). For ENCODE data, read numbers were normalized to 1x sequencing depth.

8. Statistical analysis

Box plots and violin plots were generated using R package ggplot2 (Wickham, 2009). For the box plots, the bottom and top of the box are always the first and third quartiles, and the line inside the box is always the second quartile (the median). The ends of the whiskers represent the lowest datum still within 1.5 interquartile range (IQR) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile. Any data not included between the whiskers is plotted as an outlier with a dot. In the violin plots, the median average value is represented with a dot and the vertical line represents the mean plus or minus a constant times the standard deviation (SD); the $\log_2$(FC) enrichment of the indicated ChIP-Seq signal over the input is plotted.

Bar graphs were generated with Microsoft Excel v15.33. Statistical significance was calculated with a two-tailed unpaired Student's t-test (Microsoft Excel v15.33). The data in the graphs represent the mean ± SD of independent experiments. A p-value<0.05 was considered significant (*p<0.05; **p<0.01; ***p<0.001). All experiments were performed independently at least three times.

Results

1. Characterization of DYRK1A chromatin putative binding partners

1.1. Analysis of BRCA1 and KAISO as putative DYRK1A binding partners in the context of the TCTCGCGAGA consensus motif

DYRK1A was found to be recruited mostly to proximal promoter regions enriched in the palindromic sequence TCTCGCGAGA. In addition, DYRK1A was found in protein complexes bound to this sequence *in vitro,* thereby pointing to the consensus motif as a DYRK1A binding site (Di Vona et al., 2015). To further characterize the mechanism of DYRK1A recruitment to the DNA, EMSA analysis was performed using bacterially purified GST-DYRK1A and a radiolabeled double strand probe containing the consensus motif. As a positive control, binding of bacterially expressed and purified GST-E2F1 to its consensus DNA sequence was used.

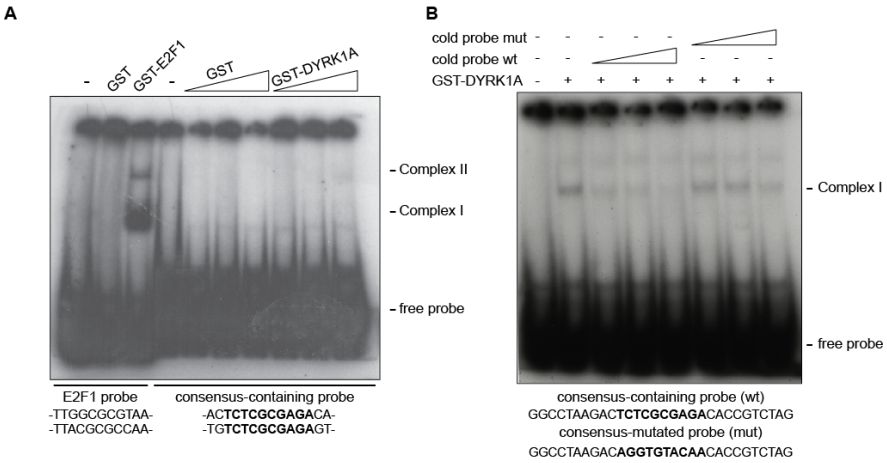


Figure R1: DYRK1A interacts in a direct manner *in vitro* with ssDNA containing the TCTCGCGAGA sequence. (A) EMSA was performed using bacterially produced GST-DYRK1A or GST-E2F1 and a ^{32}P-labeled double-stranded probe containing the DYRK1A-motif or the E2F1 binding site, respectively (the corresponding sequences are shown in the Figure). DNA-protein complexes were resolved by electrophoresis on native polyacrylamide gels and signal was assessed by autoradiography. The position of the free probe and of the shifted bands corresponding to E2F1 complexes is indicated. **(B)** A competition assay was performed by using increasing amounts (4-, 10- and 20-fold) of a single strand wild type cold probe (wt) or increasing amounts of a probe harboring mutations in the consensus sequence (mut). The position of the free probe and of the shifted band corresponding to a putative complex (complex I) is indicated.

No band-shift was detected when assaying GST-DYRK1A binding to the DYRK1A consensus-containing probe (Fig. R1A). By contrast, when a single strand DNA (ssDNA) oligonucleotide containing the palindromic

sequence was used, a retarded product was detected (Fig. R1B). Competition assays with increasing amounts of the wild type unlabeled probe blocked the formation of the complex, while incubation with increasing amounts of an oligonucleotide harboring mutations in the consensus sequence failed to do it (Fig. R1B), supporting the specificity of the direct interaction of DYRK1A with the consensus motif *in vitro*.

The DYRK1A-motif seems to be sufficient for DYRK1A recruitment to the DNA *in vitro.* Nevertheless, it is possible that other factors are required for the stabilization of DYRK1A onto DNA and/or for its function *in vivo.* To uncover putative DYRK1A binding partners in the context of the consensus sequence, analysis of ChIP-Seq data coming from the lymphoblastoid cell line GM12878 available at the ENCODE database revealed that CHD2, KAISO and BRCA1 ChIP-Seq datasets have an enrichment in this particular sequence (Di Vona, 2013; Kheradpour and Kellis, 2014; Wang et al., 2012). In addition, density plots showing the chromatin occupancies of these factors demonstrated that they overlap at the DYRK1A peak center (Fig. R2), suggesting a putative cross-talk among them. To explore this possibility, we aimed to validate the ENCODE data in the cell model used in the thesis work, the T98G cell line, for each of the factors.

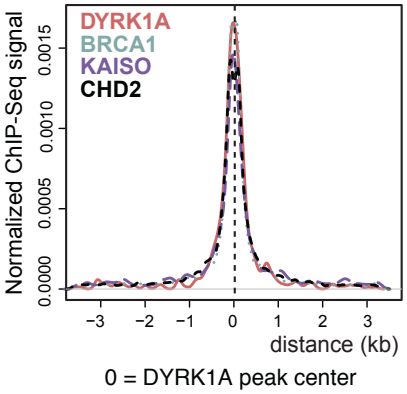


Figure R2: CHD2, BRCA1 and KAISO binding profiles relative to the DYRK1A peak center. Distribution of CHD2, BRCA1 and KAISO normalized ChIP-Seq signal across DYRK1A peak center. The plot was normalized by dividing the ChIP-Seq signal at each DYRK1A site by the maximum ChIP-Seq signal. CHD2, BRCA1 and KAISO ChIP-Seq data correspond to the GM12878 cell line obtained from the ENCODE database, while DYRK1A data corresponds to the T98G cell line (Di Vona et al., 2015). The offset was set to +/- 3 kb from the DYRK1A peak center.

1.1.1. Validation of CHD2

ChIP-qPCR with the anti-CHD2 antibody used to generate the ENCODE data (Abcam ab68301) validated the presence of CHD2 at targets associated with DYRK1A (Fig. R3B, AB). However, no CHD2 occupancy could be detected when using a different anti-CHD2 antibody (Fig. R3B, CS), which could be indicating that this antibody was not working in ChIP experiments. However, the immunoprecipitates with one of the antibodies were not detected by WB with the other antibody (Fig. R3C). As an additional control, the antibodies were tested with extracts from cells in

which the expression of CHD2 was downregulated with two independent shRNAs. Surprisingly, while the CHD2-CS antibody recognized a band of the expected molecular weight that was reduced in the silencing conditions, the CHD2-AB antibody recognized a band that was insensitive to the shRNA treatment (Fig. R3D). These observations indicate that the CHD2-AB antibody used to generate the ENCODE data is not really detecting CHD2 but an unknown protein. In fact, we aimed to identify the unknown protein by mass spectrometry in the CHD2-AB immunoprecipitates; however, the experiment was not possible because the antibody was discontinued by the company. Based on all these results, CHD2 was no longer considered in this study.

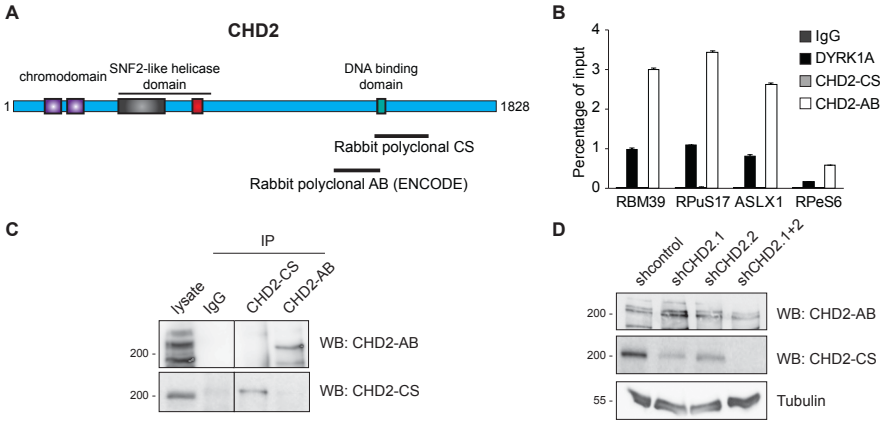


Figure R3: CHD2 is not at DYRK1A-associated genomic regions. (A) CHD2 protein structure indicating the main domains (further detailed in Fig. I5), and the epitopes used to generate two rabbit polyclonal antibodies: CS, antibody from Cell Signaling (4170) raised with a synthetic peptide corresponding to residues surrounding phenylalanine (Phe) 1579; AB, antibody from Abcam (ab68301) raised with the protein fragment 1250-1350. **(B)** ChIP-qPCR results with antibodies to DYRK1A and CHD2; normal IgGs were used as negative control (mean±SD from 3 technical replicates). **(C)** T98G soluble extracts were immunoprecipitated with the two anti-CHD2 antibodies or normal IgGs. Both lysates (10%) and immunocomplexes were analyzed by WB with the two anti-CHD2 antibodies. **(D)** CHD2 knock-down was performed in T98G cells with two independent shRNAs, either separately or co-administrated. Total extracts were analyzed by WB with the two anti-CHD2 antibodies. Tubulin was used as loading control.

1.1.2. Validation of KAISO

KAISO occupancy was validated at three DYRK1A-associated promoters containing the consensus motif, and with two independent antibodies, including the one use in ENCODE (6F8) (Fig. R4A and B). The specificity of the KAISO antibodies was evaluated by WB using extracts from cells treated with specific KAISO siRNAs or shRNAs: the anti-KAISO antibodies 558A and 6F8 detected a downregulated protein band at the expected molecular weight (Fig. R4C); however, anti-KAISO 557A did not detect a

specific band (data not shown). To further validate the specificity of the 557A antibody, immunoprecipitates obtained with this antibody were probed in WB with anti-KAISO-558A, which detected a band of the expected molecular weight (Fig. R4D). Based on these results, the use of the anti-KAISO antibodies was as follows: 6F8 for ChIP, 557A for IP and 558A for WB.

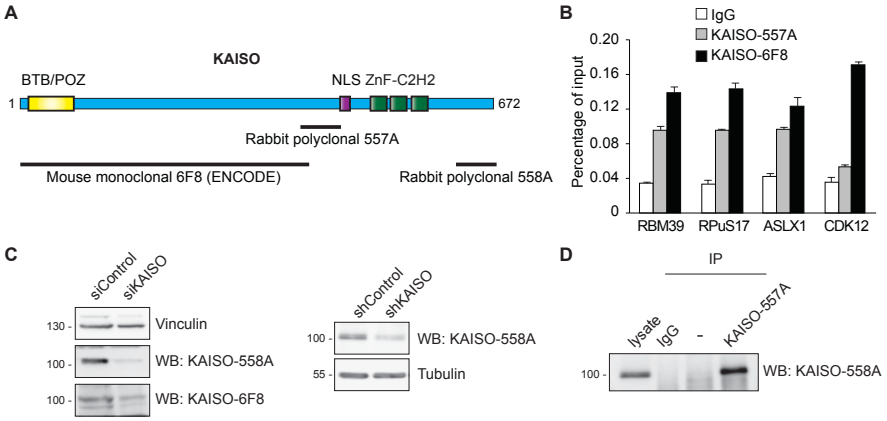


Figure R4: Validation of KAISO at some DYRK1A-associated gene targets. (A) KAISO protein structure indicating the main domains (further detailed in Fig. I6), and the epitopes used to raise three different antibodies: 6F8, a monoclonal antibody from Santa Cruz (sc-23871) raised against aa 1-504; 557A, a rabbit polyclonal antibody from Bethyl (A303-557A) raised with a synthetic peptide corresponding to aa 375-425; 558A, a rabbit polyclonal antibody from Bethyl (A303-558A) raised with a synthetic peptide corresponding to aa 622-672. **(B)** ChIP-qPCR results using anti-KAISO 557A and 6F8 antibodies and normal IgGs as negative control (mean±SD from 3 technical replicates). **(C)** KAISO expression in T98G cells was knocked down by transfection of siRNAs (left panel) or lentiviral transduction of a shRNA (right panel). Total cell extracts were analyzed by WB with the anti-KAISO antibodies indicated. Vinculin or tubulin were used as loading controls. **(D)** T98G soluble extracts were immunoprecipitated with anti-KAISO-557A or normal IgGs as negative control. Both lysates (10%) and immunocomplexes were analyzed by WB with anti-KAISO-558A antibody.

1.1.3. Validation of BRCA1

For BRCA1, four different commercial antibodies were used in the validation assays (Fig. R5A). The presence of BRCA1 at two DYRK1A motif positive promoter targets was revealed with anti-BRCA1-000A (used to generate the ENCODE data) and D20 antibodies (Fig. R5B). The antibodies were evaluated by WB of cell extracts in which BRCA1 expression was downregulated by shRNA expression: only anti-BRCA1 D9 recognized a band that was reduced in the silenced cells (Fig. R5C). In addition, anti-BRCA1 D20 and 000A antibodies were validated in IP assays (Fig. R5D). Based on these results, the BRCA1 antibodies used in this work were 000A in ChIP and D9 in IP and WB assays.

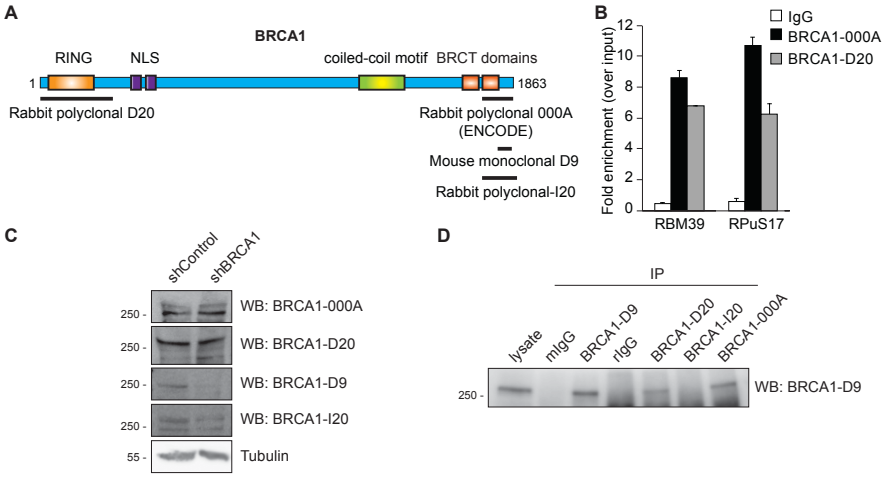


Figure R5: Validation of BRCA1 at some DYRK1A-associated gene targets. (A) BRCA1 protein structure with the main domains (further detailed in Fig. I8), and the regions used to raise the antibodies used in the study: D9, monoclonal antibody from Santa Cruz (sc-6954), raised against a peptide containing aa 1842-1862; 000A, rabbit polyclonal from Bethyl (A300-000A), raised against a peptide covering aa 1800-1863; I20, rabbit polyclonal from Santa Cruz (sc-646), raised against a peptide mapping near the C-terminus of BRCA1; and D20, rabbit polyclonal from Santa Cruz (sc-641), raised against a peptide mapping at the N-terminus. **(B)** ChIP-qPCR results using anti-BRCA1-000A and D20 antibodies and normal IgGs as negative control (mean±SD from 3 technical replicates). **(C)** BRCA1 knock-down in T98G cells was performed by lentiviral delivery of a shRNA. Total cell extracts were analyzed by WB with the anti-BRCA1 antibodies. Tubulin was used as loading control. **(D)** Hela nuclear extracts were immunoprecipitated with the anti-BRCA1 antibodies indicated or with normal mouse IgGs (mIgG) and rabbit IgGs (rIgG) as negative controls. Both lysates (10%) and immunocomplexes were analyzed by WB with anti-BRCA2-D9 antibody.

1.2. DYRK1A, BRCA1 and KAISO co-occupy promoters of genes involved in ribosome-related functions

Given the presence of BRCA1 and KAISO at several DYRK1A-associated promoters in T98G cells, we decided to carry out a genome-wide approach by ChIP-Seq experiments. Around 2,000 target regions were identified associated with DYRK1A, which meant an increment in the number of peaks compared to our previous data (Di Vona et al., 2015). This fact is likely due to technical improvements implemented at several steps of the procedure and to the gain in genome coverage thanks to the combination of several biological replicates (n=4). Interestingly, although "promoter" is still one of the categories overrepresented, the gain is more noticeable in the "intron" category, with an increment of around 10% (Fig. R6A and B; compared with Fig. I6B). GO-terms enrichment analysis showed that ribosome-associated processes and RNA processing were functional categories enriched among the DYRK1A target genes (Fig. R6C). Despite the increment in the number of DYRK1A-associated

proximal promoters and 5'-UTRs in the current ChIP-Seq dataset (from 310 to 880), GO terms analysis is in agreement to our previous results (Di Vona et al., 2015).

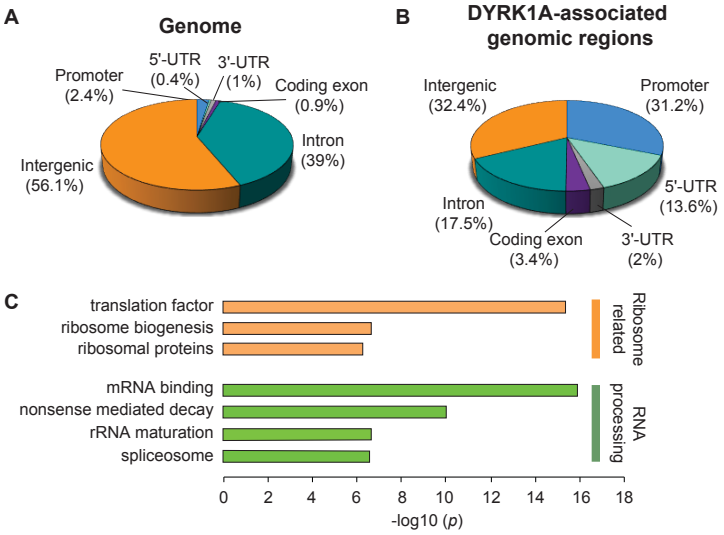


Figure R6: Characterization of DYRK1A-associated genomic *loci*. (A, B) Pie chart illustrating the general genomic distribution (**A**) or that of DYRK1A-associated targets (**B**) over some genomic features. The graph is mutually exclusive; thus, the sum of the percentage values is 100%. "Intergenic" refers to the percentage of regions that do not belong to any of the other categories. **(C)** Selected relevant enriched terms of the genes associated with DYRK1A target regions mapping to promoters, exons, introns, 5'- and 3'-UTRs.

For KAISO and BRCA1, we obtained around 500 and 300 target regions, respectively. In both cases, there was a clear enrichment of promoters over the rest of genomic features (Fig. R7A and B).

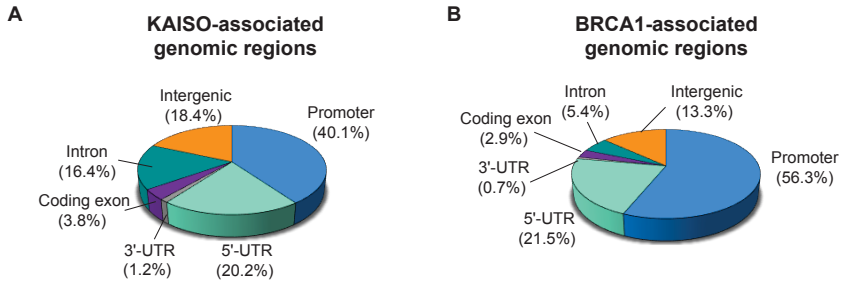


Figure R7: KAISO and BRCA1 are located mostly at promoter regions. Pie chart illustrating the distribution of KAISO (**A**) and BRCA1 (**B**) -associated targets over some genomic features.

Next, we aimed to characterize the genomic regions co-occupied by DYRK1A, KAISO and BRCA1. First, we wanted to confirm the existence

of a pool of genomic regions recognized by the three factors in the cell system used in this thesis work, the T98G cell line. As shown in Fig. R8, BRCA1 and KAISO occupy genomic regions overlapping with DYRK1A-associated peaks. This finding not only confirms the previous results obtained using the ENCODE data sets (Fig. R2), but also suggests that the co-localization of DYRK1A with KAISO and BRCA1 at genomic *loci* might be common to different cell types.

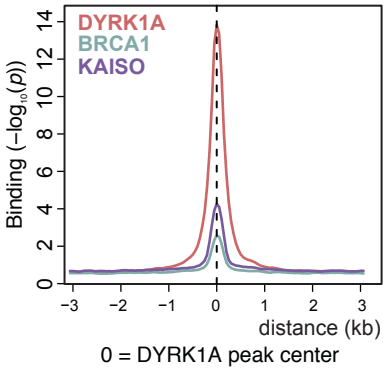


Figure R8: DYRK1A, KAISO and BRCA1 binding profiles at DYRK1A-associated genomic regions in T98G cells. Density plots showing the distribution of DYRK1A, KAISO and BRCA1 relative to the DYRK1A peak center. The offset was set to +/- 3 kb from the DYRK1A peak center. The y-axis represents relative protein recruitment calculated as the $-\log_{10}$ (*p*-value) of the ChIP-Seq data over the input.

To identify distinctive features, we decided to characterize the most different subsets of genes: on one hand, the ones occupied by the three factors and, on the other hand, the ones exclusively occupied by each of them. Analysis of the ChIP-Seq data revealed that the genomic regions co-occupied by the three proteins represent a small subset of the DYRK1A ChIP genomic *loci* (Fig. R9A: ≈ 7% of total DYRK1A ChIP peaks). However, the average density of reads is significantly higher in the pool of co-occupied ChIP targets compared to that corresponding to the regions bound only by DYRK1A (Fig. R9B and C), indicating a stronger presence of DYRK1A at the sites co-bound by the three factors.

Annotation of the *cis*-regulatory elements present in each of the clusters showed that the co-occupied genomic *loci* correspond mostly to promoters, whereas introns and intergenic regions are overrepresented in those DYRK1A-specific sites (Fig. R9D). In addition, motif enrichment analysis found the DYRK1A-motif as strongly enriched in the co-occupied targets (Fig. R9E). Of note, MEME catalogues this motif as a KAISO binding site. The larger group of DYRK1A-specific regions also showed this particular motif (Fig. R9E). However, the enrichment was highly significant for motifs corresponding to CTCF binding sites (Fig. R9E). GO terms enrichment analysis revealed that ribosome-related processes are overrepresented in the subset of target genes co-bound by the three factors (Fig. R9F), whereas RNA processing-associated functions are associated with gene targets belonging to both clusters (Fig. R9F and G).

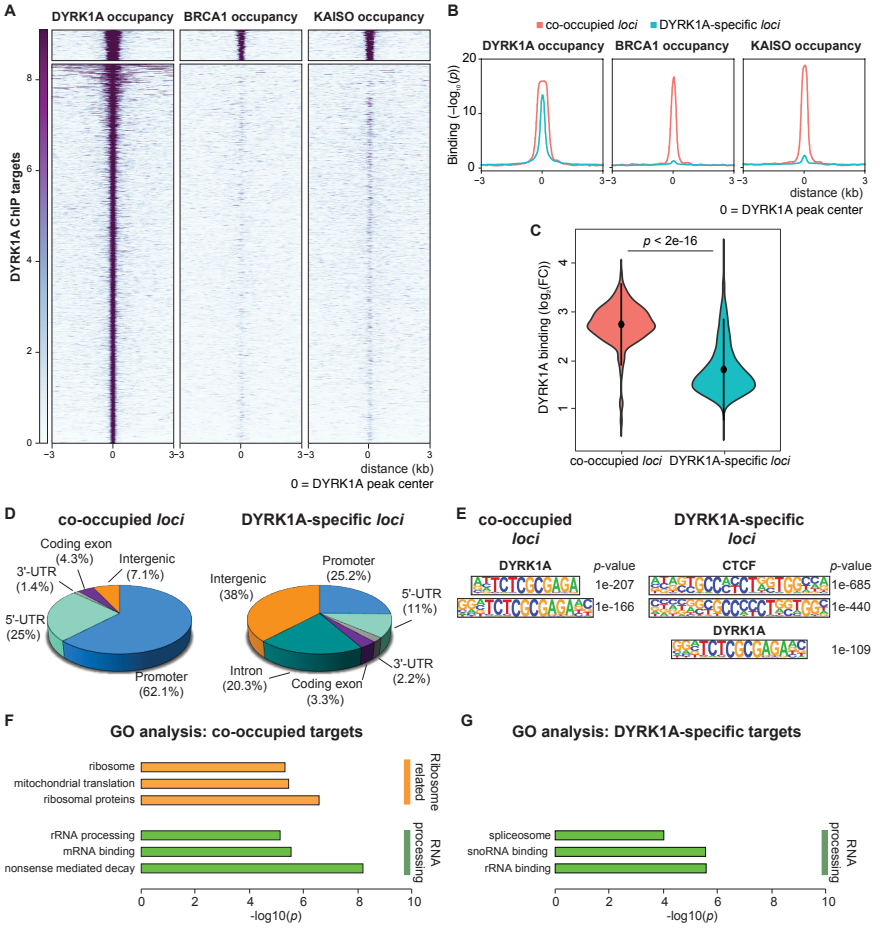


Figure R9: The genomic regions co-occupied by DYRK1A, BRCA1 and KAISO display distinctive features. (A) Heatmaps showing DYRK1A, BRCA1 and KAISO occupancies for DYRK1A ChIP regions relative to the DYRK1A peak center. The color bar scale indicates the binding score calculated as described in MM5.5. The ChIP-associated regions were separated in two clusters to differentiate co-occupied regions *vs* the DYRK1A-specific ChIP targets. **(B)** Density plots indicating the average binding following the same criteria as in (**A**). **(C)** Violin plot showing the DYRK1A ChIP signal in co-occupied regions and in DYRK1A-specific regions; p<2e-16, Student's t-test. **(D)** Pie chart illustrating the distribution over some genomic features of DYRK1A-BRCA1-KAISO co-occupied targets and of targets occupied only by DYRK1A. **(E)** Motif enrichment analysis of co-occupied targets and of DYRK1A-only target regions using MEME. **(F, G)** Selected relevant enriched terms of DYRK1A-BRCA1-KAISO (**F**) and of DYRK1A-specific (**G**) of gene targets (include ChIP regions mapping to promoters, exons, introns, 5'-UTRs and 3'-UTRs).

The same analysis was performed using the KAISO or BRCA1 ChIP datasets as reference (Fig. R10 and R11). In the case of KAISO, approximately 27% of the KAISO ChIP peaks were occupied by the three factors. Similar to the case of DYRK1A, the average density of reads of the KAISO ChIP targets co-occupied by DYRK1A and BRCA1 is significantly higher than the one of the KAISO-specific subset, thus

indicating that KAISO is more represented in the first category (Fig. R10A-C). In addition, the co-bound regions are enriched in promoters and in the DYRK1A-consensus motif, while the rest of the KAISO-associated genomic *loci* are mostly introns and intergenic regions, not particularly enriched in any TF binding site (Fig. R10D and E). Interestingly, GO terms analysis showed no significantly overrepresented functions for the genes associated with KAISO-only occupation.

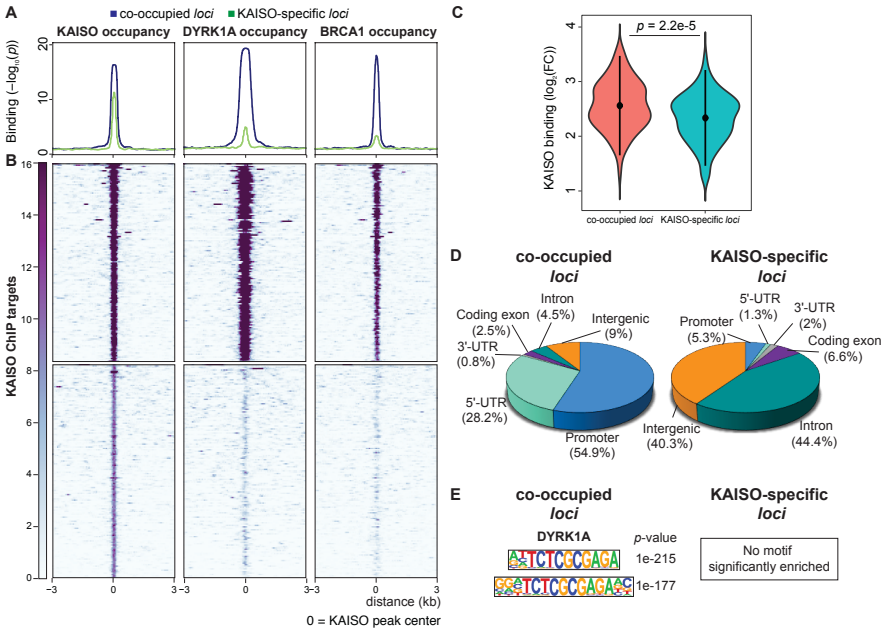


Figure R10: The KAISO genomic regions co-occupied by DYRK1A and BRCA1 are mostly promoters enriched in the DYRK1A-motif. (A) Density plots showing KAISO, DYRK1A and BRCA1 occupancies for KAISO ChIP regions relative to the KAISO peak center. The binding score was calculated as described in MM5.5. The ChIP-associated regions were separated in two clusters to differentiate co-occupied regions *vs* the KAISO-specific ChIP targets. **(B)** Heatmaps indicating the ChIP average binding of KAISO, DYRK1A and BRCA1 for BRCA1 ChIP regions following the same criteria as in (**A**). The color bar scale indicates the binding score calculated as described in MM5.5. **(C)** Violin plot showing the KAISO ChIP signal in co-occupied regions and in KAISO-specific regions; *p*=2.2e-5, Student's *t*-test. **(D)** Pie charts illustrating the distribution over some genomic features of DYRK1A-BRCA1-KAISO co-occupied target regions or the KAISO-specific *loci*. **(E)** Motif enrichment analysis by MEME in the two clusters. The search rendered no motif particularly enriched in the KAISO-specific set of target regions.

The analysis of the BRCA1 dataset showed that approximately 47% of total BRCA1 ChIP targets are occupied by both DYRK1A and KAISO, BRCA1 ChIP binding is similar in both the co-bound and BRCA1-specific gene subsets (Fig. R11A-C). As expected, the analysis further confirmed that promoters and the DYRK1A-motif are features overrepresented in the subset of genomic regions occupied by the three factors (Fig. R11D and

E). Although the other cluster of BRCA1-associated *loci* contains a considerable number of promoters, they do not present any enrichment in particular TF binding sites (Fig. R11D and E). Furthermore, GO terms analysis showed that there is no function overrepresented among the gene targets belonging to this category. In conclusion, these observations suggest a possible interplay among DYRK1A, KAISO and BRCA1 at the promoter level of genes mainly related to ribosome functions.

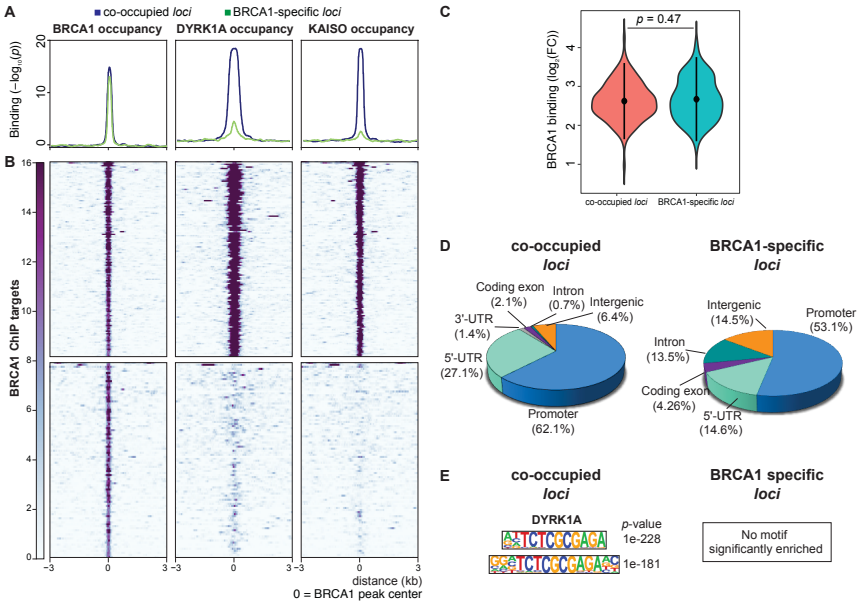


Figure R11. The BRCA1 genomic regions co-occupied by DYRK1A and KAISO are mostly promoters enriched in the DYRK1A-motif. (A) Density plots showing BRCA1, DYRK1A and KAISO occupancies for BRCA1 ChIP regions relative to the BRCA1 peak center. The binding score was calculated as described in MM5.5. The ChIP-associated regions were separated in two clusters to differentiate co-occupied regions *vs* the BRCA1-specific ChIP targets. **(B)** Heatmaps indicating the ChIP average binding of BRCA1, DYRK1A and KAISO for BRCA1 ChIP regions following the same criteria as in (**A**). The color bar scale indicates the binding score calculated as described in MM5.5. **(C)** Violin plot showing the BRCA1 ChIP signal in co-occupied regions and in BRCA1-specific regions; p=0.47, Student's t-test. **(D)** Pie charts illustrating the distribution over some genomic features of DYRK1A-BRCA1-KAISO co-occupied target regions or BRCA1-specific *loci*. **(E)** Motif enrichment analysis by MEME in the two clusters. The search rendered no motif particularly enriched in the BRCA1-specific set of target regions.

1.3. KAISO and DYRK1A do not have a clear cross-talk at the chromatin level

Next, we aimed to investigate the putative cross-talk between DYRK1A and the two factors. First, we wanted to know where KAISO localizes in T98G cells, and a subcellular fractionation procedure was carried out. The results showed that KAISO was mostly found in the soluble nuclear

compartment (Fig. R12A). By contrast, DYRK1A was mostly found in the cytoplasm, which is in agreement with previous data showing the subcellular distribution of DYRK1A in HeLa cells (Di Vona et al., 2015). A pool of both proteins was also observed in the chromatin insoluble fraction (Fig. R12A), likely responsible for their presence at chromatin.

To assess the DYRK1A and KAISO relationship, we asked whether these two proteins interact. Immunoprecipitation assays with T98G total soluble extracts showed that DYRK1A was found in complexes immunoprecipitated with an anti-KAISO antibody, whereas an antibody against DYRK1A failed to co-immunoprecipitate KAISO (Fig. R12B). A plausible interpretation of this result is that the more abundant DYRK1A cytosolic pool does not interact with KAISO, while nuclear KAISO is able to interact with nuclear DYRK1A. To test this hypothesis, immunoprecipitation experiments were carried out using concentrated commercial HeLa nuclear extracts with similar results (Fig. R12C). These findings could be interpreted as result of transient interactions, which cannot be captured with the DYRK1A antibodies. In addition, we cannot rule out that the interaction does happen on chromatin and/or that it depends on specific signals.

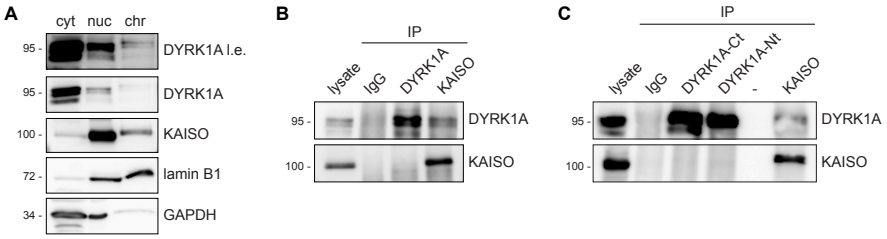


Figure R12: DYRK1A and KAISO might be part of protein complexes in the nucleus. (A) Subcellular fractionation of T98G cells was performed to separate into cytosolic (cyt), nuclear soluble (nuc) and nuclear-insoluble/chromatin (chr) fractions. The samples were analyzed by WB with antibodies for the indicated proteins. Lamin B1 and GAPDH were used as nuclear and cytosolic markers, respectively. For DYRK1A, a longer exposure (l.e.) is also included. **(B)** T98G soluble extracts were immunoprecipitated with anti-DYRK1A or anti-KAISO antibodies or normal IgGs as negative control. The lysate (10%) and immunocomplexes were analyzed by WB with antibodies to the indicated proteins. **(C)** HeLa nuclear extracts were immunoprecipitated with two antibodies anti-DYRK1A, either targeting the C-terminal part (DYRK1A-Ct) or the N-terminus of DYRK1A (DYRK1A-Nt), or with an anti-KAISO antibody. Normal IgGs were used as negative controls. The lysate (10%) and immunocomplexes were analyzed by WB with antibodies to the indicated proteins.

Despite the low level of interaction, DYRK1A could be a kinase for KAISO. This hypothesis was tested in radioactive *IVK* assays with bacterially-produced GST-KAISO. However, no incorporation of radiolabeled ATP was observed, which demonstrated that KAISO was not a DYRK1A substrate (Fig. R13).

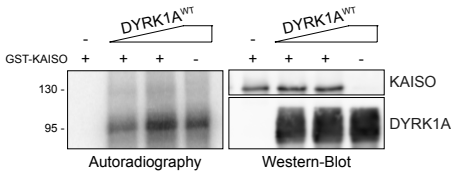


Figure R13: DYRK1A does not phosphorylate KAISO. Bacterially produced GST-KAISO was subjected to an *IVK* assay in the presence of DYRK1A and [γ^{32}P]-ATP. Samples were resolved by SDS-PAGE and incorporation of ^{32}P analyzed by autoradiography. The presence of the indicated proteins was assessed by WB with a specific antibody for DYRK1A and a GST antibody for KAISO. Note that autophosphorylation of DYRK1A is detected.

To elucidate a possible cross-talk between DYRK1A and KAISO at the chromatin level, KAISO occupancy was analyzed upon DYRK1A depletion by lentiviral delivery of a shRNA to DYRK1A and *vice versa* (Fig. R14 and R15). ChIP-Seq experiments showed that DYRK1A expression reduction was also reflected at its associated genomic *loci* (Fig. R14B). However, no major loss of KAISO occupancy at KAISO-associated sites was observed under the same conditions (Fig. R14C).

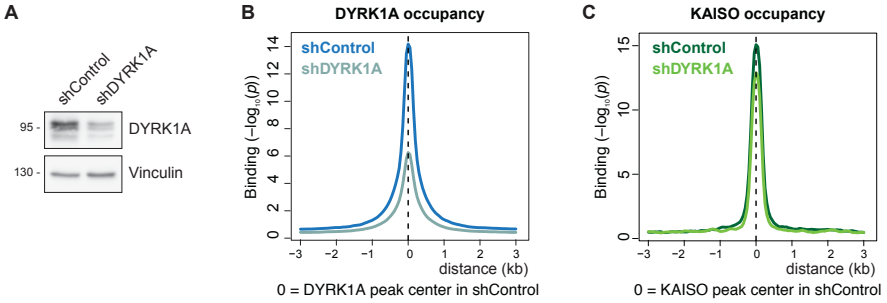


Figure R14: KAISO chromatin occupancy is not globally affected upon DYRK1A silencing. (A) The reduction in DYRK1A levels in the shDYRK1A-T98G cells used in the ChIP assays was assessed by WB. An antibody to vinculin was used as loading control. **(B)** Density plots were generated to show the distribution of DYRK1A in shControl and shDYRK1A conditions relative to DYRK1A peak center in shControl cells. **(C)** Density plots showing the distribution of KAISO in shControl and shDYRK1A conditions relative to KAISO peak center in shControl cells. The offset was set to +/- 3 kb.

Likewise, no apparent changes in DYRK1A occupancy were observed in cells depleted of KAISO (Fig. R15B). At the time of writing this report, no ChIP data for KAISO in conditions of KAISO silencing have been obtained. Therefore, we cannot rule out that the pool of KAISO associated with chromatin is not affected by the silencing conditions, so that these results will need to be re-evaluated when the data is available.

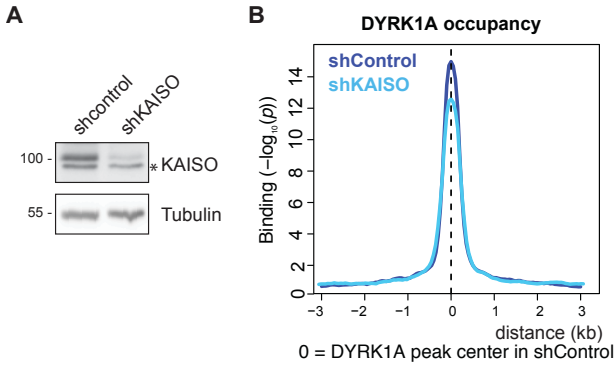


Figure R15: Chromatin-bound DYRK1A is not globally altered in KAISO-depleted cells. **(A)** The reduction in KAISO levels in the shKAISO-T98G cells used in the ChIP assays was assessed by WB. An antibody to tubulin was used as loading control. *, non-specific band. **(B)** Density plots showing the distribution of DYRK1A in shControl and shKAISO conditions relative to the DYRK1A peak center in shControl cells. The offset was set to +/- 3 kb.

Note that, for both KAISO and DYRK1A, chromatin occupancies tend to decrease when the expression of the other factor is reduced (Fig. R14C and R15B). This observation might suggest that the recruitment of each protein could depend on the other at specific ChIP targets. However, this possibility has not been addressed in the frame of this thesis work.

1.4. BRCA1 does not affect DYRK1A chromatin recruitment

To identify whether BRCA1 is involved in the recruitment of DYRK1A to its genomic *loci*, we analyzed DYRK1A chromatin occupancy by ChIP-Seq analysis in BRCA1-depleted cells. Given the fact that BRCA1 is a E3-ubiquitin ligase able to target substrates for proteasome degradation, we first checked whether DYRK1A protein levels were altered and found no effect in the absence of BRCA1 (Fig. R16A). These results do not exclude the possibility of DYRK1A being a target of BRCA1 in non-degradative ubiquitination.

Reduction of BRCA1 expression led to a reduction of around 50% in BRCA1 binding to its targets (Fig. R16B). By contrast, no loss of DYRK1A chromatin binding at DYRK1A ChIP regions was observed under BRCA1-silencing conditions (Fig. R16C).

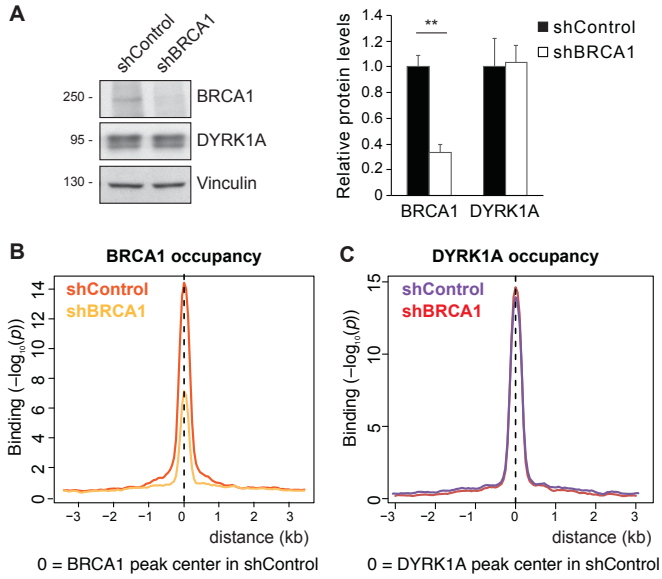


Figure R16: BRCA1 is not involved in DYRK1A chromatin recruitment. (A) Total protein levels of DYRK1A and BRCA1 were analyzed by WB in T98G cells infected with lentivirus expressing shBRCA1 or shControl. Vinculin was used as loading control. A representative image is shown. The graph shows the quantification of the protein bands relative to vinculin, with the shControl levels arbitrarily set as 1 (mean±SD of 3 independent experiments; Student's *t*-test, $^{**}p \leq 0.01$). **(B)** Density plots indicating the distribution of BRCA1 at chromatin in shControl and shBRCA1 T98G cells relative to BRCA1 peak center in shControl conditions. **(C)** Density plots indicating the distribution of DYRK1A at chromatin in shControl and shBRCA1 T98G cells relative to DYRK1A peak center in shControl conditions. The offset was set to +/- 3 kb.

1.5. DYRK1A is involved in BRCA1 chromatin recruitment

Next, we asked whether DYRK1A plays a role in BRCA1 chromatin recruitment. To answer this question, ChIP-Seq experiments assessing BRCA1 occupancy upon DYRK1A depletion were performed. The results indicated that no reorganization of BRCA1 to other genomic regions was observed (data not shown), but BRCA1 occupancy at its genomic targets was significantly reduced in shDYRK1A T98G cells compared to shControl ones (Fig. R17A). The reduction in BRCA1 occupancy was observed both at regions recognized by DYRK1A and at those ones where only BRCA1 is present (Fig. R17B and C; see representative images in Fig. R17D). For further validation, two BRCA1-DYRK1A target regions (*RPS11/uS17* and *RPS15/uS19*) and one target occupied only by BRCA1 (*MRPL16*) were selected to perform ChIP-qPCR experiments. As shown in Fig. R17E, DYRK1A was found enriched in the expected targets and responded to the knockdown. On the other hand, BRCA1 presence was reduced upon DYRK1A silencing at the co-occupied targets, while no such effect was observed in the *MRPL16* specific case (Fig. R17E).

Therefore, while the results validated the ChIP-Seq data on the co-occupied promoters, further experiments would be needed to confirm the results on the BRCA1-unique sites.

Based on these results, it is possible that DYRK1A has a direct effect on BRCA1 chromatin eviction at the common sites. However, we cannot rule out that DYRK1A could also regulate BRCA1 chromatin-associated functions outside the chromatin.

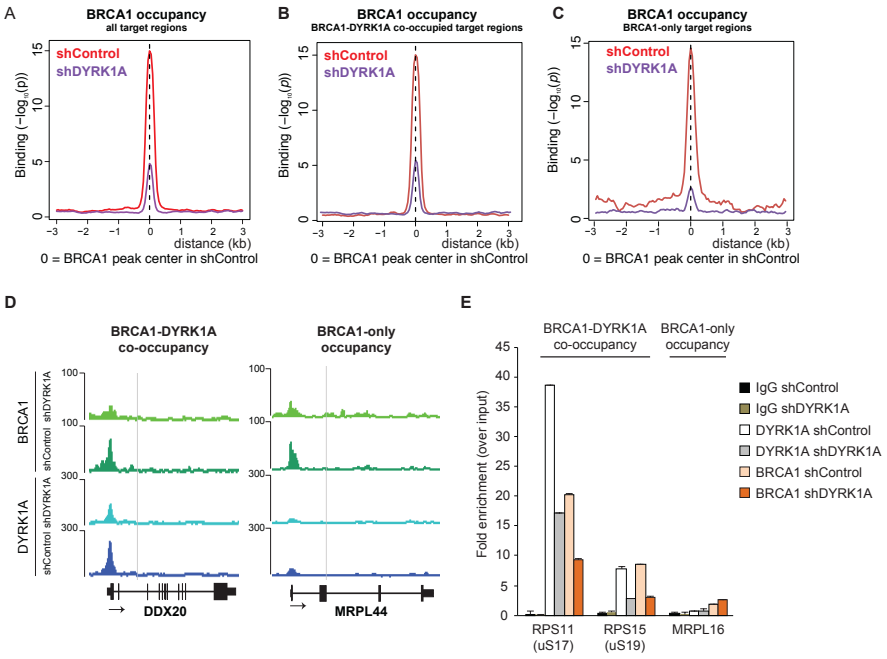


Figure R17: DYRK1A downregulation has an impact in BRCA1 chromatin occupancy. **(A-C)** Density plots indicating the distribution of BRCA1 at total BRCA1 ChIP targets (**A**), DYRK1A-BRCA1 co-occupied regions (**B**), and BRCA1 unique genomic regions (**C**) in shControl or shDYRK1A T98G cells relative to BRCA1 peak center in shControl conditions. **(D)** Representative examples of DYRK1A and BRCA1 occupancies in shControl and shDYRK1A-infected conditions at proximal promoters of one BRCA1-DYRK1A co-occupied gene target (left) and a BRCA1 unique ChIP target (right). The arrows indicate direction of transcription. **(E)** Validation of selected targets by ChIP-qPCR in shControl or shDYRK1A T98G cells using anti-BRCA1, anti-DYRK1A antibodies or IgGs as negative control. The experiment was performed with external "spike-in" as described in MM5.2 and 5.3. Data is represented as DNA recovery over the input (mean±SD of 3 technical replicates).

To gain further insight into the underlying mechanism in the DYRK1A-BRCA1 cross-talk, we first determined whether BRCA1 and DYRK1A interact. BRCA1 is described to locate mostly in the nucleus in several cell types such as MCF7 or HeLa cell lines (Hernandez et al., 2018; Nepomuceno et al., 2017), so we started by analyzing the subcellular localization on BRCA1 in T98G cells. Subcellular fractionation

experiments showed that BRCA1 mostly accumulates in the soluble nuclear fraction (Fig. R18A). Based on this result, the immunoprecipitation assays were performed using commercial Hela nuclear extracts. BRCA1 was found in anti-DYRK1A complexes obtained with three different antibodies (Fig. R18B). In addition, DYRK1A was co-immunoprecipitated with an antibody anti-BRCA1 (Fig. R18C). These results indicate that DYRK1A and BRCA1 are part of common complexes, at least in the nucleus.

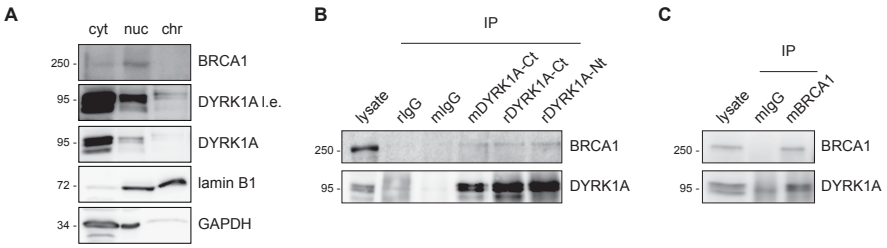


Figure R18: DYRK1A and BRCA1 interact in the nucleus. (A) Subcellular fractionation was performed to separate into cytosolic (cyt), nuclear soluble (nuc) and nuclear-insoluble/chromatin (chr) fractions. The samples were analyzed by WB with antibodies for the indicated proteins. Lamin B1 and GAPDH were used as nuclear and cytosolic markers, respectively. For DYRK1A, a longer exposure (l.e.) is also included. **(B, C)** Hela nuclear extracts were immunoprecipitated with three anti-DYRK1A antibodies (mDYRK1A, mouse monoclonal; rDYRK1A-Ct, rabbit polyclonal targeting the C-terminal part; rDYRK1A-Nt, rabbit polyclonal targeting the N-terminus) (**B**) or an anti-BRCA1 antibody (mBRCA1, mouse monoclonal) (**C**). Normal rabbit or mouse IgGs (rIgG or mIgG) were used as negative controls. The lysate (10%) and the immunocomplexes were analyzed by WB with antibodies to the indicated proteins.

Next, we wondered whether DYRK1A downregulation has an impact in either the expression of BRCA1 or its subcellular localization. Results from WB experiments in DYRK1A-knocked down T98G cells indicated that BRCA1 total protein levels were not affected under these conditions (Fig. R19A). Interestingly, preliminary subcellular fractionation data showed a subtle reduction of BRCA1 in the insoluble nuclear fraction when DYRK1A is silenced (Fig. R19B), a mechanism that would be in agreement with BRCA1 eviction from its genomic targets.

Since DYRK1A is a kinase, we wondered whether DYRK1A-dependent effects on BRCA1 chromatin occupancy could be triggered by phosphorylating BRCA1. To test this hypothesis, radioactive *IVK* assays were performed on overexpressed Flag-tagged BRCA1 that was affinity purified from soluble cell extracts using an anti-Flag antibody. A signal over the background corresponding to the molecular size of Flag-BRCA1 could be detected only in the immunocomplexes incubated with DYRK1A wild-type, but not with a kinase-inactive mutant (Fig. R20), indicating that BRCA1 is a DYRK1A substrate, at least, *in vitro*.

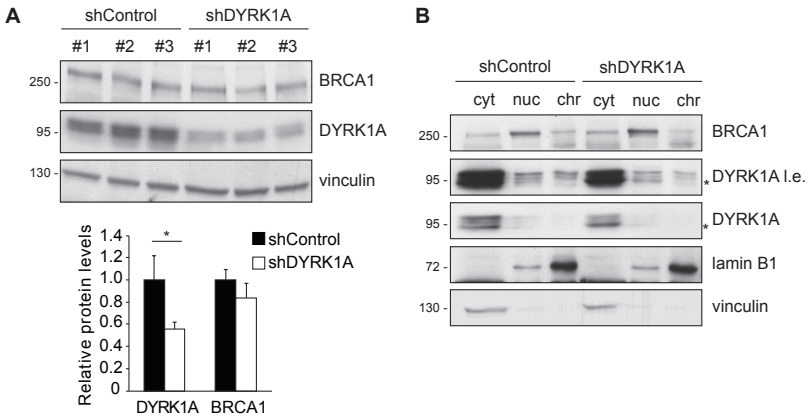


Figure R19: DYRK1A downregulation leads to BRCA1 relocation from nucleus to the cytoplasm. (A) Upper panel: Total protein levels of DYRK1A and BRCA1 were analyzed by WB in T98G cells infected with lentivirus expressing shDYRK1A or shControl. Vinculin was used as loading control. **Lower panel:** The graph shows the quantification of the protein bands relative to vinculin, with the shControl levels arbitrarily set as 1 (mean±SD of 3 independent experiments); Student's *t*-test, *p≤0.05. **(B)** Subcellular fractionation was performed to separate into cytosolic (cyt), nuclear soluble (nuc) and nuclear-insoluble/chromatin (chr) fractions in shControl and shDYRK1A T98G cells. Fractions were analyzed by WB with antibodies for the indicated proteins. For DYRK1A, a longer exposure (l.e.) is shown. Lamin B1 and vinculin were used as markers for the nuclear and cytosolic fraction, respectively. *, non-specific band.

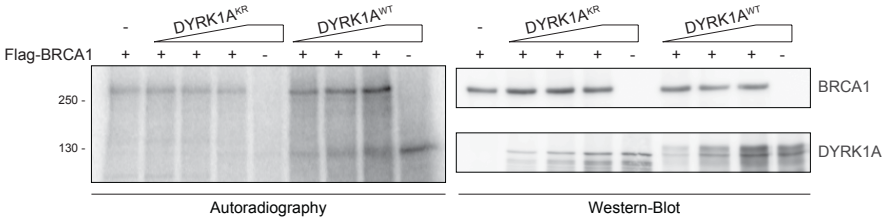


Figure R20: BRCA1 is phosphorylated by DYRK1A. *IVK* assays using Flag-tagged BRCA1 expressed in HEK-293T cells and immunopurified with an anti-Flag antibody were incubated with purified GST-DYRK1A wild-type ($DYRK1A^{WT}$) or a kinase-death version ($DYRK1A^{KR}$) in the presence of ^{32}P-γ-ATP. Proteins were separated by SDS-PAGE and analyzed by autoradiography. The presence of the proteins was confirmed by WB (BRCA1, anti-Flag; DYRK1A, anti-GST). Background BRCA1 phosphorylation was detected in the Flag-BRCA1 immunocomplexes likely due to the effect of an accompanying kinase.

To identify the phosphorylated residues, and given that BRCA1 is a big protein to be expressed in bacteria, we decided to fragment the protein in products of 300 aa (see MM2.2 for the cloning strategy), and expressed them in bacteria as GST-fusion proteins (Fig. R21A). Each of the GST-BRCA1 fragments was used in *IVK* assays with purified GST-DYRK1A and radiolabeled ATP. The summary of the experiments is shown in Fig.

R21A, which shows that all BRCA1 fragments, except the fragment containing the first 300 aa (N), were phosphorylated by DYRK1A. These results indicate that DYRK1A phosphorylates BRCA1 at several residues along the primary sequence.

The identification of the phosphosites was done by MS of the purified fragments phosphorylated *in vitro*. The analysis rendered a peptide coverage of 70% and several phosphopeptides were identified scattered along BRCA1 protein sequence (Fig. R21B). Interestingly, phosphorylation events in several BRCA1 residues have been described to regulate BRCA1 subcellular localization and chromatin recruitment (Altiok et al., 1999; Kehn et al., 2007), having a further impact on BRCA1-dependent transcriptional mechanisms (Hinton et al., 2007). Therefore, the phosphorylation of BRCA1 by DYRK1A could be playing a role in this context and it will be further addressed in the future.

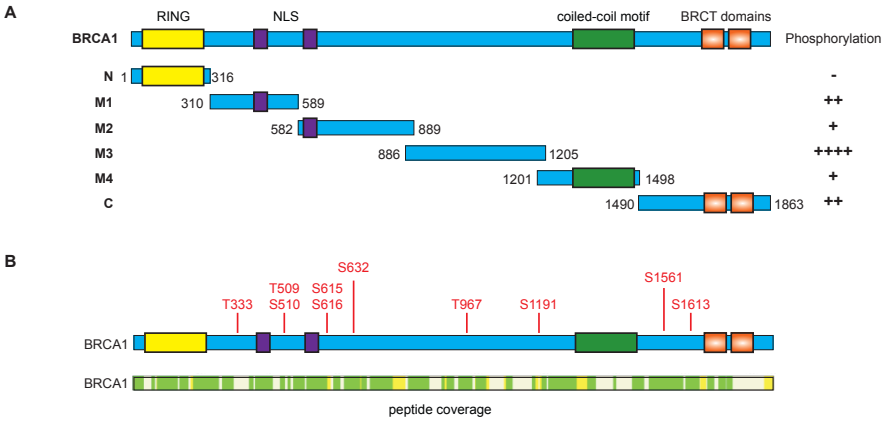


Figure R21: DYRK1A phosphorylates BRCA1 at multiple sites. (A) Hot *IVK* assays were performed on GST-BRCA1 fragments covering the coding sequence of the protein: amino-terminal fragment N (1-316 aa); middle fragments M1 (310-589 aa), M2 (582-889 aa), M3 (886-1205 aa), and M4 (1201-1498 aa); and carboxy-terminal fragment C (1490-1863 aa). A summary of the results is shown with the symbol + indicating the intensity of the radioactive signal. **(B)** Cold *IVK* assays were performed on each of the GST-BRCA1 fragments and the phosphorylated peptides were identified by MS analysis. The position of the phosphosites is indicated in red; peptide coverage is also shown (FDR≤1%, green; FDR≤5%, yellow).

1.6. DYRK1A-dependent BRCA1 chromatin eviction does not have a clear impact on gene expression

Since both DYRK1A and BRCA1 have been described to play roles in transcription, we wondered whether there is a correlation between their chromatin occupancies and alterations in the mRNA levels of the corresponding targets when their expression is reduced. To assess putative transcriptional roles at the chromatin level, we intersected RNA-

Seq data with ChIP-Seq data. Thinking in a direct transcriptional effect, only peaks mapping to promoters has to be considered; however, DYRK1A and BRCA1 ChIP peaks associated with promoters are very close to the TSS, making these regions difficult to be accurately distinguished from 5'-UTRs during peak annotation. Therefore, we considered both genomic features, promoters and 5'-UTRs, for the analysis.

RNA-Seq experiments in shDYRK1A T98G cells compared to control ones showed that 2,794 genes were downregulated whereas 609 genes were upregulated with a fold change cutoff of +/- 1.5 (Fig. R22A). The overlap with the DYRK1A ChIP dataset indicated that 191 out of 880 DYRK1A target regions mapping to promoters and 5'-UTRs showed altered transcript levels in their associated genes when DYRK1A expression was reduced. Interestingly, the majority of them, 185 out of 191, turned to be downregulated (Fig. R22B), thus supporting the role of DYRK1A as a transcriptional activator.

Downregulation of BRCA1 also induced alterations in transcription profiles with 2,070 genes upregulated and 1,739 genes downregulated with a fold change cutoff of +/- 1.5 (Fig. R22C). In contrast to the DYRK1A results, the association of BRCA1 chromatin recruitment with transcriptional regulation was not so clear: although a similar proportion of BRCA1 target regions mapping to promoters and 5'-UTRs present alterations in the expression of associated genes upon BRCA1 silencing (73 out of 282), some of the genes were downregulated (27 genes) while others were found upregulated (46 genes) (Fig. R22D).

To gain further insight into the putative impact of BRCA1-DYRK1A cross-talk in transcription, we focused on the pattern of expression presented by the DYRK1A-BRCA1 co-bound ChIP targets. Taking advantage of the intersection between the RNA-Seq and ChIP-Seq data, we observed that 62 DYRK1A-BRCA1 co-occupied gene targets were downregulated in DYRK1A-reduced conditions, but only 16 of them showed alterations in their mRNA levels in BRCA1-knocked down cells compared to control ones, which were not consistent in the direction of change (Fig. R22B). Complementarily, despite the majority of BRCA1-associated regions whose expression is altered upon BRCA1 depletion are also recognized by DYRK1A (67 out of 73), only 16 of them were downregulated in a DYRK1A-dependent manner (Fig. R22D).

These observations indicate that the loss of BRCA1 at its genomic *loci* does not have a common impact in the expression of DYRK1A-BRCA1 co-occupied gene targets. Therefore, the consequences of the DYRK1A-dependent BRCA1 chromatin eviction have to be clarified.

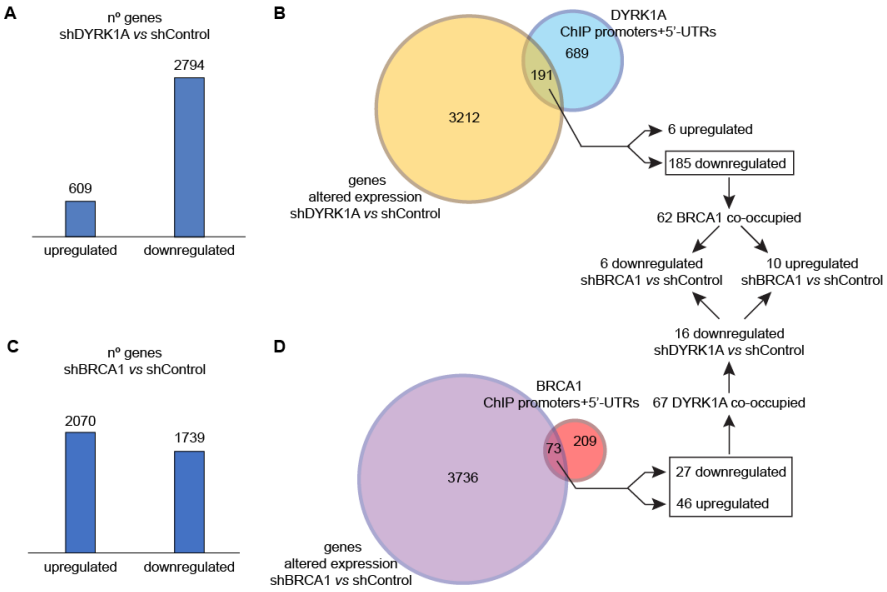


Figure 22: The proportion of DYRK1A-BRCA1 co-occupied gene targets that present concordant alterations in expression is low. (A) Bar graph showing the number of genes with altered expression in shDYRK1A T98G cells. The RNA-Seq data was normalized with an external spike-in as explained in MM6.2 and 6.5 (-1.5 < FC > 1.5; $p \leq 0.05$). **(C)** Bar graph showing the number of genes with altered expression in shBRCA1 T98G cells compared to shControl ones (-1.5 < FC > 1.5; $p \leq 0.05$). **(B, D)** Venn diagram indicating the overlap between DYRK1A ChIP targets mapping to proximal promoters and 5'-UTRs and genes whose transcript levels are altered upon DYRK1A silencing (**B**) or between BRCA1 ChIP targets mapping to proximal promoters and 5'-UTRs and genes whose transcript levels are altered upon BRCA1 silencing (**D**). The number of gene targets co-occupied by both factors is indicated in each dataset as well as of those found altered upon silencing of the other factor.

2. Role of chromatin-bound DYRK1A on the transcriptional regulation of ribosomal proteins

2.1. Chromatin-bound DYRK1A as a transcriptional activator of ribosome-related genes

The overlap between the DYRK1A ChIP-Seq data and the RNA-Seq data in conditions of DYRK1A silencing indicates that DYRK1A may regulate the transcriptional activation of at least 185 gene targets (Fig. R23A). GO terms analysis in this particular subset of genes showed that they are involved in functions related to ribosome processes, RNA processing and DNA damage (Fig. R23B). To further validate these results, we assessed the mRNA expression of specific gene targets belonging to the three categories by RT-qPCR: the ribosome-related category was represented by three RPGs, *RPS19/eS19, RPS6/eS6* and *RPL17/uL22*; the RNA processing category was represented by *UTP6* (U3 small nucleolar RNA-

associated protein 6), *RBM28* (RNA binding motif protein 28) and *NOL11* (nucleolar protein 11); and in representation of the last category, we analyzed the mRNA levels of the replication factors *RFC3* and *RFC5* and of the DNA polymerase ε subunit, *POLE*. All these gene targets showed a significant reduction in their mRNA levels in shDYRK1A T98G cells compared to control ones (Fig. R23C). These results confirm that DYRK1A acts as a transcriptional activator in a subset of ChIP target genes involved in ribosome-related processes, RNA processing and DNA replication.

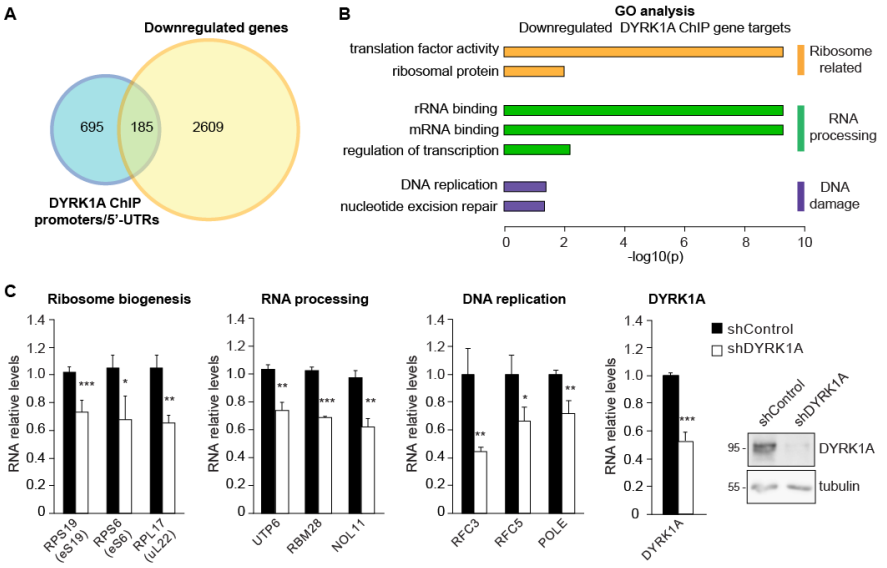
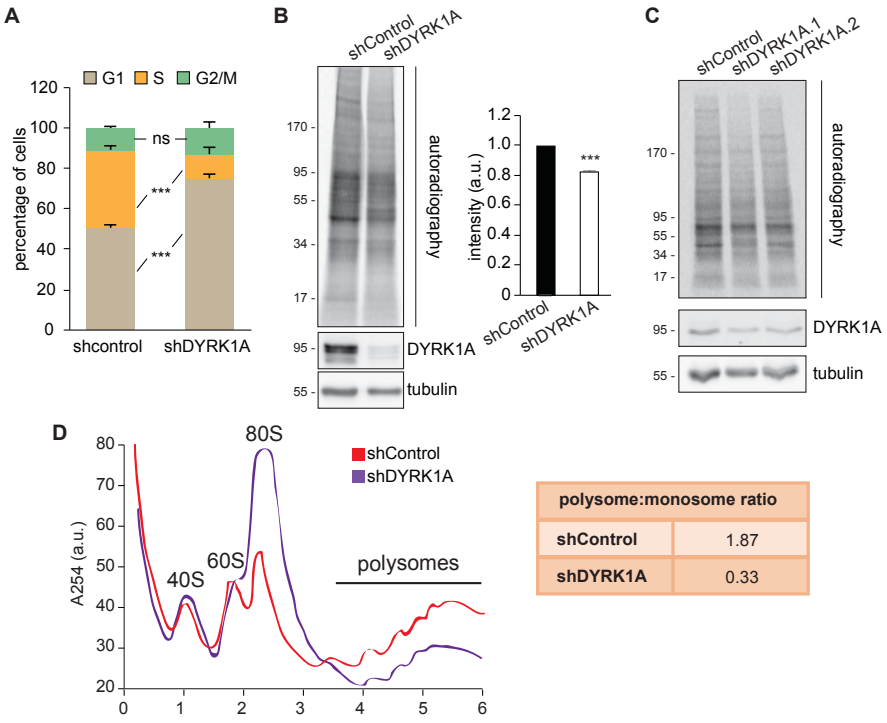


Figure R23: A subset of DYRK1A ChIP targets is downregulated upon DYRK1A depletion. (A) Venn diagram indicating the overlap between DYRK1A ChIP target genes mapping to proximal promoters and 5'-UTRs and those genes found to be downregulated upon DYRK1A silencing in the RNA-Seq experiment (p<0.05; -1.5< FC >1.5) **(B)** GO terms analysis of the genes overlapped in panel **A**. **(C)** Validation by RT-qPCR of the expression of selected DYRK1A targets in each of the categories in T98G cells transduced with shControl or shDYRK1A lentiviruses. The experiment was performed using an external spike-in as described in MM6.2 and 6.3. Data are represented as RNA levels relative to the shControl condition set as 1 (mean±SD, n=3 biological replicates; Student's *t*-test, ***$p\leq0.001$, **$p\leq0.01$, *$p\leq0.05$). The reduction in DYRK1A expression was shown by RT-qPCR (mean±SD, n=3 biological replicates; Student's *t*-test, ***$p\leq0.001$) and by WB of one of the samples.

2.2. DYRK1A-depleted cells present a phenotype defective in protein translation

Given the fact that one of the categories mostly enriched in DYRK1A ChIP targets downregulated in a DYRK1A-dependent manner is related to ribosome functionality, we wondered whether DYRK1A could be contributing, at least in part, to cell growth by affecting the transcription of key proteins. Concurring with this idea, previous work from our laboratory has shown that T98G cells present a reduction in cell volume upon DYRK1A downregulation (Di Vona, 2013). This effect has been observed in many other cell lines (data not shown), suggesting that it is a common phenotypic output of DYRK1A downregulation. The alteration in cell size could be due to a defect in cell cycle and/or to a loss in protein content.

The first possibility was evaluated by assessing the cell cycle profile of DYRK1A-downregulated T98G cells. The results indicated that there is an increase of cells in G1 and a decrease of cells in S when DYRK1A expression is reduced (Fig. R24A). The alternative possibility was addressed with radiolabeled ^{35}S-methionine incorporation assays. The results revealed that DYRK1A-depleted cells with two independent shRNAs present a reduction in their translational rate (Fig. R24B and C). These observations could reflect a complex scenario where DYRK1A-dependent cell cycle and cell size regulators could be contributing to the final phenotype. Additionally, the functional status of ribosomes was analyzed by polysome profiling upon DYRK1A silencing. This technique allows observing to what extent ribosomes are engaged in active translation, generally considered as such when they are organized in polysomes. In agreement with the reduction in translation rate, shDYRK1A-T98G cells presented a decrease in the polysome fraction compensated with an increase in the monosome peak compared to control ones (Fig. R24D). These results indicate that DYRK1A depletion leads to polysome disorganization so that DYRK1A should be involved somehow in regulating protein synthesis mechanisms.



polysome:monosome ratio	
shControl	1.87
shDYRK1A	0.33

Figure R24: DYRK1A-depletion leads to alterations in protein synthesis. (A) Cell cycle profile of cells infected with lentiviruses expressing a shRNA control or shDYRK1A.1. The graph shows the percentage of cells in each cell cycle phase (mean±SD of 3 biological replicates; Student's *t*-test, ***$p \leq 0.001$, ns=not significant). **(B, C)** ^{35}S-labeled methionine incorporation assays were performed in shDYRK1A or shControl T98G cells. The incorporation of radioactive Met was evaluated by autoradiography. Protein loading and the presence of DYRK1A were analyzed by WB. A representative experiment and the quantification of average intensity of three independent experiments with shDYRK1A.1 (**B**) or a single experiment with two independent shRNAs (**C**) are shown. Student's *t*-test, ***$p \leq 0.001$. **(D)** Polysome profiles of DYRK1A-downregulated T98G cells (purple line) compared to control cells (red line). The positions of the 40S, 60S, 80S, and polysomal peaks are indicated. The *y*-axis shows absorbance at 254 nm in arbitrary units and the *x*-axis relative fractions. The polysome-monosome ratio is shown for each condition as the ratio between the areas under the curve of the 80S peak *vs* the polysome-peaks.

We are aware that the effects of DYRK1A downregulation on translation rates could be indirect by acting on signaling pathways relevant to this process. Previous data from our laboratory demonstrated that phosphorylation of ribosomal protein S6 kinase, a final event of the mTOR signaling pathway, does not change in DYRK1A-silenced T98G cells (Di Vona, 2013). Despite this result, the alteration could be still due to DYRK1A downregulation inducing cellular stress, which by converging in the phosphorylation of the translation initiation factor eIF2α, activates pathways leading to protein synthesis inhibition (Koromilas, 2015).

However, the levels of eIF2α phosphorylation remained unchanged in the absence of the kinase (Fig. R25), thus excluding this possibility.

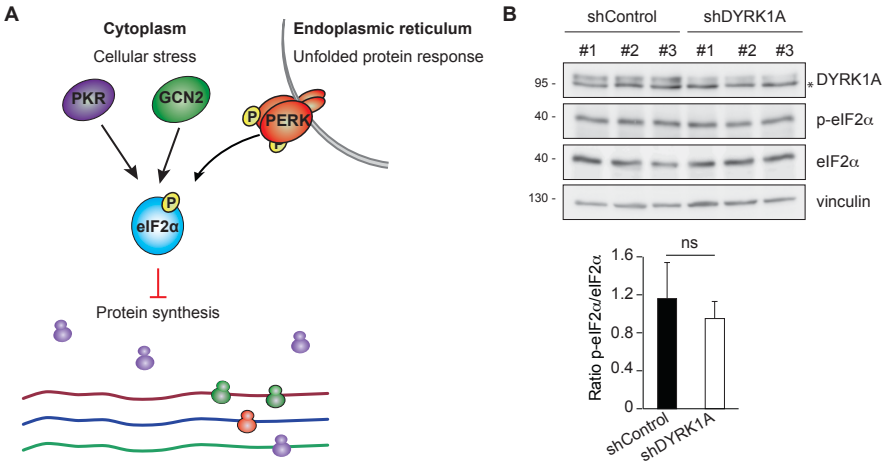


Figure R25: DYRK1A downregulation does not alter eIF2α phosphorylation levels. (A) Cartoon depicting the key factors involved in cellular stress and unfolded protein response pathways leading to eIF2α phosphorylation and reduced protein synthesis rate: PERK, protein kinase RNA-like endoplasmic reticulum kinase; PKR, protein kinase double-stranded RNA-dependent; GCN2, general control non-derepressible 2. **(B)** Analysis by WB of the levels of eIF2α phosphorylation in Ser51 in T98G control or shDYRK1A-cell samples. *, non-specific band. The graph shows the quantification of the signal pS51-eIF2α to total eIF2α (mean±SD of 3 independent experiments; Student's *t*-test, ns= not significant).

Although we cannot rule out the potential impact of DYRK1A on other signaling pathways converging on translation regulation, altogether the results support the hypothesis of DYRK1A playing a role in protein synthesis downstream in the cascade, probably at the promoter level of key factors involved in the process.

2.3. Chromatin-bound DYRK1A as a transcriptional regulator of canonical ribosomal proteins

As shown in Fig. R26A, the kinase seats very close to the TSS in around 25% of the RPGs. Interestingly, DYRK1A-associated RPG promoters are characterized by the presence of a highly conserved DYRK1A-motif (*q*-value<0.05), whereas those RPG promoters lacking DYRK1A do not contain this motif or present degenerated versions of the sequence (*q*-value>0.06). Of note, the kinase was found in a minority of RPG promoters without the palindromic consensus, thereby pointing out to a motif-independent way of DYRK1A recruitment to a particular set of RPGs (Fig. R26B for some examples of the different categories). It is therefore

possible that DYRK1A acts as a transcriptional regulator of a subset of RPGs directly at their promoter level.

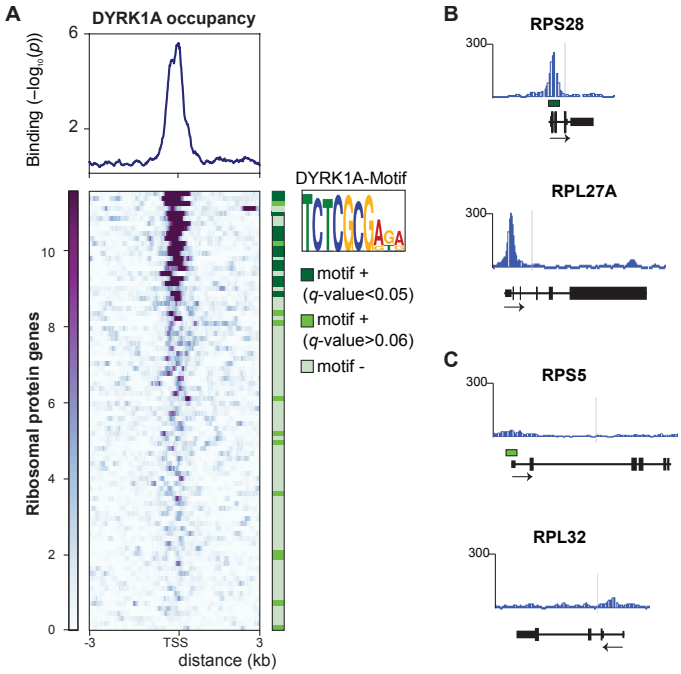


Figure R26: DYRK1A is overrepresented at RPG promoters containing the DYRK1A-consensus. (A) Density plot indicating DYRK1A ChIP binding relative to the TSS of all RPGs. The heatmap shows DYRK1A occupancy across all RPGs relative to the TSS. The binding score is indicated in the density plot and in the color bar scale, and it was calculated as described in MM5.5. The position of the TSSs was taken from the UCSC genome browser (SwitchGear TSS Track). The offset was set to a +/- 3 kb from the TSS. The presence of the DYRK1A-motif in each gene is indicated in dark green (*q*-value<0.05) and light green (*q*-value>0.06). **(B, C)** Representative examples of RPG promoters with (**B**) or without (**C**) DYRK1A recruitment and with or without the presence of the DYRK1A-motif. The arrows indicate the direction of transcription.

To test this hypothesis, we assessed the expression level of those RPGs that are DYRK1A ChIP targets in DYRK1A-silenced conditions in comparison to those RPGs in which DYRK1A is not detected at their promoters. Interestingly, genes belonging to the first subset tend to have higher expression levels than the ones where DYRK1A is absent in shControl-infected cells. This difference seems to be diluted upon DYRK1A depletion, thus suggesting that the enhancement in expression is DYRK1A-dependent (Fig. R27A). In addition, only those genes whose promoters are occupied by DYRK1A are significantly downregulated in shDYRK1A infected cells compared to control ones (*p*-value = 0.003) (Fig. R27A). Selected target RPGs were validated by RT-qPCR, with a reduction of expression more apparent in those genes containing

DYRK1A at their promoters (Fig. R27B, *RPL21*'s differences were not significant when 3 biological replicates were analyzed). These results suggest that DYRK1A is contributing, at least partially, to protein synthesis regulation by activating the transcription of canonical RPs directly at the promoter level.

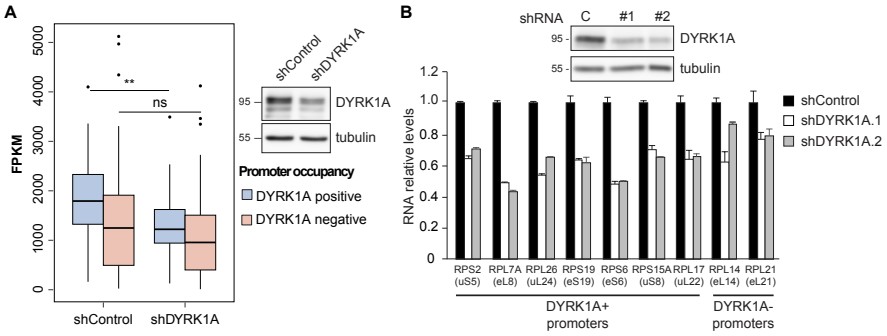


Figure R27: The expression of RPGs containing DYRK1A at promoters is significantly reduced upon DYRK1A downregulation. (A) Box-plot indicating the expression of RPGs (FPKM) whose promoters present (blue) or lack (pink) DYRK1A, both in shDYRK1A and shControl T98G cells. Student's *t*-test, ***p*=0.003, ns= not significant. The reduction in DYRK1A levels is shown by WB analysis. **(B)** Validation by RT-qPCR of selected RPGs differentially expressed in T98G with DYRK1A downregulation by lentiviral transduction of two distinct shRNAs for DYRK1A. Data are represented as RNA levels relative to the shControl condition set as 1 (mean±SD of 3 technical replicates). The reduction in DYRK1A levels is shown by WB analysis.

2.4. Characterization of DYRK1A-positive ribosomal protein gene promoters

As explained in the Introduction (I3), very little is known about the transcriptional regulation of RPGs in mammals. Given the overrepresentation of DYRK1A at these promoters and its role as a transcriptional activator, we intended to search for differential features that may characterize DYRK1A-positive RPG promoters.

First, we assessed BRCA1 and KAISO presence at RPG promoters. The analysis showed that their occupancies overlap with that of DYRK1A at the majority of motif-positive RPG promoters (Fig. R28). Unfortunately, to date, we have not observed a clear interplay between DYRK1A and KAISO at the chromatin level (Fig. R13-R15); furthermore, BRCA1 functionality at ChIP targets co-occupied with DYRK1A is yet to be elucidated in the cell model used (Fig. R16 and R22).

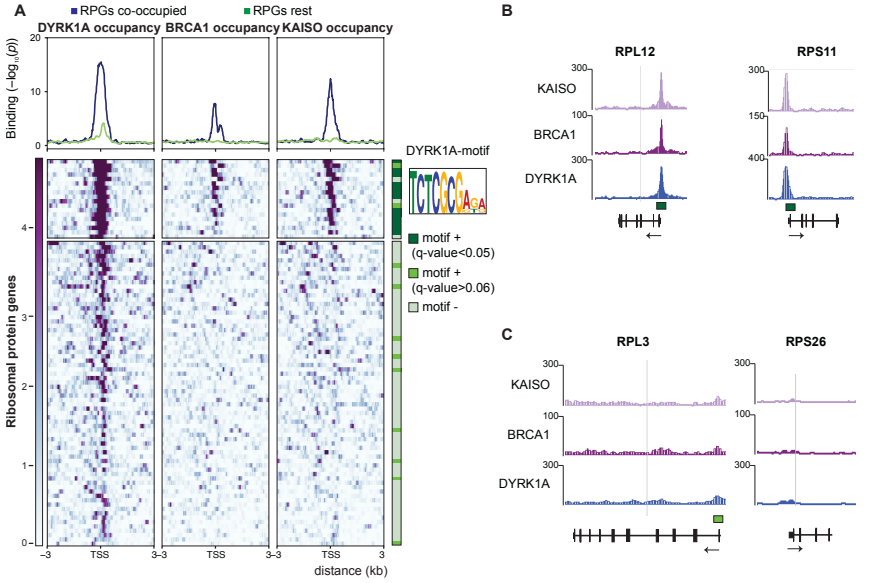


Figure R28: BRCA1 and KAISO co-occupy DYRK1A-positive RPG promoters containing highly conserved TCTCGCGAGA-like sequences. (A) Density plots indicating DYRK1A, BRCA1 and KAISO ChIP binding relative to the TSS of all RPGs. The heatmap shows DYRK1A, BRCA1 and KAISO occupancies across all RPGs relative to the TSS. The binding score is indicated in the density plot and in the color bar scale, and it was calculated as described in MM5.5. *K*-means clustering was applied resulting in co-occupied regions *vs* the rest of RPGs. Positions of the TSS were taken from the UCSC genome browser (SwitchGear TSS Track). The offset was set to a +/- 3 kb from the TSS. The presence of the DYRK1A motif in each gene is indicated in dark green (*q*-value<0.05) and light green (*q*-value>0.06). **(B, C)** Representative examples of RPGs occupied by DYRK1A, BRCA1 and KAISO (**B**) or not occupied by these factors (**C**). The presence of the DYRK1A consensus is indicated in each RPG. The arrows indicate the direction of transcription.

In the light of these results, we focused our attention in the search of alternative motif-binders. In a characterization study of RPG promoters, the palindromic consensus motif was described as a hDRE-like sequence, potentially recognized by hDREF (Yamashita et al., 2007) (further detailed in I3.2). We therefore checked whether hDREF is present at DYRK1A promoter targets containing the palindromic sequence. However, we were unable to detect hDREF at several of those sites (Fig. R29A). In addition, the two proteins were not found in common complexes in soluble nuclear extracts (Fig. R29B). These results indicate that hDREF and DYRK1A might not be common regulators of RPGs expression. It should be mentioned that although the authors conclude that the DYRK1A-motif is similar to a previously described DRE binding motif in *Drosophila*, we think that the similarity between the two motifs is very low, which would explain why the binding is not detected.

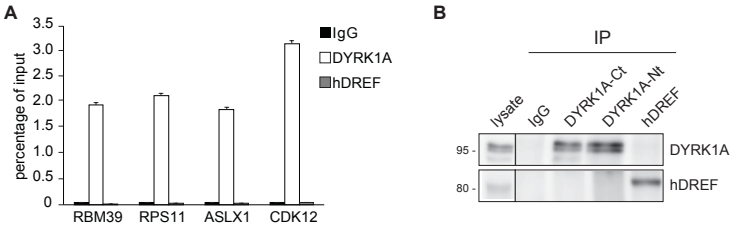


Figure R29: hDREF is not a chromatin binding partner of DYRK1A. (A) ChIP-qPCR experiments using DYRK1A or hDREF antibodies and normal IgGs as negative control. The data is represented as percentage of input recovery (mean ± SD of three technical replicates). **(B)** HeLa nuclear extracts were immunoprecipitated with two antibodies anti-DYRK1A, either targeting the C-terminal part (DYRK1A-Ct) or the N-terminus of DYRK1A (DYRK1A-Nt), or anti-hDREF specific antibodies. Normal IgGs were used as negative control. The lysate (10%) and the immunocomplexes were analyzed by WB.

To search for further DYRK1A co-regulators of RPG expression, we investigate whether other TF binding sites were differentially overrepresented in those RPG promoters containing the DYRK1A-motif. To approach this issue, we performed an unbiased MEME search of binding sites in the subset of RPGs catalogued as motif-positive. Interestingly, the analysis showed that the binding site associated with the orphan Nuclear receptor subfamily 4 group A member 1 (NR4A1) appeared in 60% of the genes belonging to this subset (Fig. R30A). On the other hand, the RPG promoters lacking the DYRK1A consensus showed an enrichment in ETS binding sites (Fig. R30B). Therefore, there is a differential distribution of TF binding sites according to the presence or absence of the DYRK1A-consensus motif.

Given the fact that the chromatin recruitment of the factors is not always constrained to the presence of their canonical binding sites, we asked whether the observed differential distribution of motifs correlates with differential chromatin occupancy of their corresponding factors. To approach this question, we analyzed available public ChIP-Seq data of selected candidates. Interestingly, the presence of binding sites for one ETS family member, GABP, has been associated with RPG promoters (Perina et al., 2011; Perry, 2005) as well as its occupancy at *RPL30/eL30* and *RPL32/eL32* promoters (Genuario et al., 1993; Yoganathan et al., 1992). In the light of these observations, we performed *k*-means clustering on DYRK1A, NR4A1 and GABP occupancy at RPG promoters using ChIP-Seq data obtained from T98G cells (this thesis work), Kasumi leukemia cells (Duren et al., 2016), and in SK-N-SH neuroblastoma cells (ENCODE), respectively. We also included YY1 in the study (SK-N-SH cells, ENCODE) because its binding site is enriched in RPG promoters (Perina et al., 2011; Perry, 2005) and because YY1 interacts with DYRK1A when exogenously expressed (Salichs, 2008). Of note, GABP

and YY1 ChIP data could be obtained from a cell line closely related to the cell model used in the Thesis, and indeed a positive correlation analysis of the KAISO chromatin profile in T98G and SK-N-SH was indicative of conserved interactions (data not shown).

The analysis demonstrated that NR4A1 is homogeneously found across all the clusters with poor occupancy, invalidating the motif-enrichment data for this TF (Fig. R30A and C). Likewise, YY1 is not preferentially located at DYRK1A-occupied RPG promoters (Fig. R30C). By contrast, DYRK1A and GABP were differentially distributed: cluster 1 presented promoters depleted of DYRK1A but highly occupied by GABP; cluster 2 is represented by RPG promoters with high DYRK1A occupancy and low GABP presence; and cluster 3 is characterized by promoters without DYRK1A and low levels of GABP (Fig. R30C). Indeed, the DYRK1A-motif is mostly overrepresented in cluster 2 whereas cluster 1 and 3 mainly comprise promoters with no motif or with degenerated versions (Fig. R30D). These results not only reinforce the previous data showing a high correlation between the presence of the palindromic consensus and DYRK1A at RPG promoters (Fig. R26), but also indicate that low GABP promoter occupancy is an additional feature of the DYRK1A-bound RPG promoters.

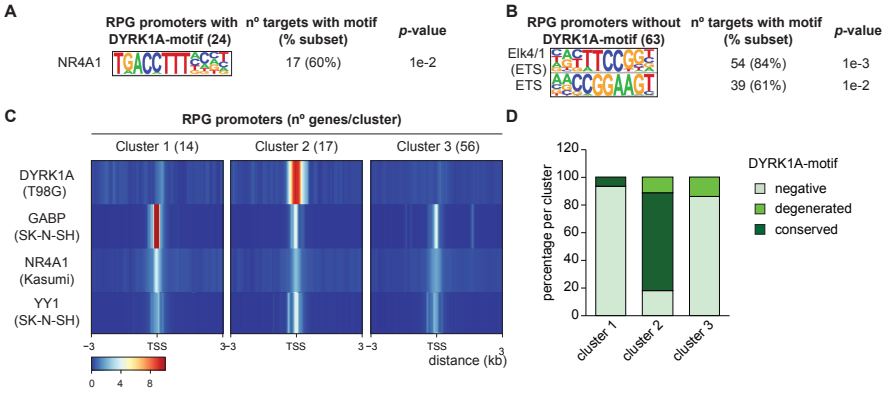


Figure 30. DYRK1A-positive RPG promoters are characterized by the presence of a NR4A1 binding site and a low GABP promoter occupancy. (A, B) Overrepresented motifs found as result of a MEME motif search in the DYRK1A-motif positive RPG cluster (**A**) *vs* the RPG group lacking the motif (**B**). The representation of each motif in both clusters and their respective *p*-values are indicated. **(C)** Average DYRK1A, GABP, NR4A1 and YY1 ChIP binding relative to the TSS of all RPGs distributed in three clusters by K-means clusterization. The cell line used for the analysis of each factor occupancy is indicated in brackets. The color bar scale indicates score of binding calculated as described in MM5.5. **(D)** The histogram represents the percentage of RPG promoters containing the DYRK1A-motif in each cluster, either conserved (dark green; *q*-value<0.05) or degenerated (light green; *q*-value>0.06).

Discussion

1. DYRK1A interacts with single strand DNA containing the consensus motif

Previous work from the laboratory demonstrated the presence of DYRK1A on complexes able to bind the DYRK1A-motif (Di Vona et al., 2015). However, no indication on whether DYRK1A was able to directly bind DNA was obtained. In this Thesis work, results from EMSA analysis using bacterially-expressed and purified DYRK1A show that DYRK1A binds the palindromic sequence TCTCGCGAGA in a direct manner *in vitro.* Interestingly, the interaction only takes place when assaying ssDNA oligonucleotides. Former results from the group had already shown that DYRK1A is able to interact with random sequence ssDNA *in vitro* (Salichs, 2008). Therefore, the new results could indicate that the sequence in the motif confers specificity to the DYRK1A interaction or/and increases the binding affinity; in any case, we cannot exclude that additional factors may be required in *vivo* to either increase the specificity and/or the strength of the interaction.

Given the fact that the DYRK1A-motif is frequently located very close to the TSS, one might speculate that DYRK1A could bind to ssDNA upon the melting of the region as a consequence of the transcription bubble opening. Therefore, it could be interesting to perform ChIP-Seq experiments to assess DYRK1A recruitment in a situation where RNAPII cannot unwind DNA, such as by inhibition of the TFIIH helicase activity (Titov et al., 2011). If true, such results would imply that DYRK1A is recruited after PIC assembly. Alternatively, DYRK1A may be directly recruited to the motif under specific circumstances where ssDNA is exposed, such as DNA damage or DNA replication. Nevertheless, considering that the presence of DYRK1A on chromatin is detected in normal growing conditions, the existence of additional mechanisms to explain DYRK1A recruitment to the consensus would be required.

2. An updated view of the DYRK1A-associated genomic regions

Technical improvements and the use of several replicates of DYRK1A ChIP-Seq data have rendered a global gain in the number of high confidence DYRK1A-associated genomic regions compared to the previous report of the group (Di Vona et al., 2015). Interestingly, the increment is more apparent within the intron category. So far, phosphorylation of the CTD at Ser5 has been linked mainly to transcription initiation, which is in agreement with DYRK1A location at

promoters and its capability to phosphorylate this residue *in vivo* (Di Vona et al., 2015). However, recent native elongating transcript sequencing (NET-Seq) data assessing nascent mRNA associated with particular RNAPII CTD modifications demonstrated that high levels of phosphoSer5 on the CTD are found at exon-intron junctions and are needed for co-transcriptional splicing to occur (Harlen et al., 2016; Nojima et al., 2015). The presence of DYRK1A at introns could be related to this co-transcriptional activity. In fact, DYRK1A has been described to phosphorylate components of the splicing machinery (Fig. I4), thus reinforcing this hypothesis.

Another interesting information coming from the DYRK1A ChIP data analysis is the overrepresentation of CTCF binding sites in *loci* depleted of BRCA1 and KAISO (Fig. R9). Preliminary analysis showed that CTCF-conserved motifs (q-value<0.05) are mostly located at DYRK1A-associated introns and intergenic regions (data not shown). CTCF is known as an insulator located, together with cohesin, mainly at the boundaries of topologically associated domains, genomic regions characterized by long-range DNA contacts (Merkenschlager and Nora, 2016). The fact that DYRK1A is located at CTCF binding sites leads to consider the involvement of DYRK1A in long-range DNA contacts as a new mechanism for the kinase to regulate transcription or the possibility for DYRK1A to act as a chromatin barrier. In addition, CTCF is also present in gene bodies and promoters (Ruiz-Velasco et al., 2017). In this regard, we have also found that a proportion of DYRK1A-associated 5'-UTRs are positive for CTCF binding sites (degenerated versions with q-value>0.06). In these cases, the presence of CTCF could serve to connect promoters and enhancers (Hanssen et al., 2017), or to act as a barrier to antisense transcription (Bornelov et al., 2015). A better understanding of the functional connection of DYRK1A and CTCF should come from results of experimental approaches that provide information on high-order chromatin interactions.

3. Potential interplay of DYRK1A and KAISO at chromatin in alternative contexts

Our results show that KAISO and DYRK1A chromatin occupancies overlap in a subset of genomic regions where promoters and the DYRK1A-motif appeared as overrepresented features (Fig. R9 and R10). Previously, we have observed that DYRK1A chromatin occupancy was partially reduced at specific targets upon KAISO silencing (Di Vona, 2013). However, the wide-genome analysis performed in this work did not show this effect at the global level (Fig.

R15); likewise, KAISO chromatin occupancy did not appear to be affected by DYRK1A reduction (Fig. R14), suggesting that the recruitment of each factor to genomic targets is independent of the presence of the other factor. Despite these results, we cannot exclude the possibility that DYRK1A and KAISO exert roles in the transcription of their common targets in a synergistic or antagonistic manner without any effect on their respective recruitments. The transcriptional activity of KAISO depending on the DYRK1A-motif is still not well understood. For instance, KAISO has been shown to recruit SMRT to the methylated palindromic sequence and cooperate with SMRT and NCoR in the repression of the expression of genes involved in terminal adipogenesis (Raghav et al., 2012). In this context, an inhibitor of class I DYRKs has been shown to abrogate adipocyte differentiation (Masaki et al., 2015). It is thus possible that DYRK1A could antagonize KAISO in the transcriptional regulation of target genes required for this process. In addition, it has been observed that KAISO interacts with TCF4 and that both factors co-occupy DYRK1A-motif-positive regions in the context of oligodendrocyte maturation. The interaction leads to a blockade of the Wnt/β-catenin pathway likely due to KAISO displacing β-catenin from TCF4 (Zhao et al., 2016). Notably, DYRK1A is able to interact with TCF4 in a β-catenin-dependent manner (Arato, 2010). Importantly, DYRK1A has been demonstrated to participate in neural cell lineage specification (further detailed in I1), thus suggesting possible antagonistic functions of KAISO and DYRK1A in this scenario. Nevertheless, the motif-positive gene promoters in adipocyte or oligodendrocyte differentiation were occupied neither by KAISO nor DYRK1A in T98G cells. Therefore, it could be interesting to assess the DYRK1A-KAISO relationship at promoters in additional cell types such as the ones described.

Finally, KAISO has been described to bind and negatively regulate CTCF at specific gene promoters (reviewed in I2.3). In addition, analysis of ENCODE data revealed that CTCF binding sites are enriched in KAISO-associated genomic *loci* in K562 cells (Blattler et al., 2013). However, our results show that CTCF binding sites are mostly enriched in DYRK1A-associated regions depleted of both KAISO and BRCA1 (Fig. R9), so it is not possible to determine whether CTCF is somehow involved in the transcription regulation of a subset of DYRK1A-KAISO promoters in a cell type-specific manner.

In summary, despite all these observations, the results obtained so far do not allow to establish any interplay between DYRK1A and KAISO at

the chromatin level, and their possible cross-talk has to be elucidated yet.

4. BRCA1 chromatin ejection depends on DYRK1A

The data collected in this Thesis work indicate that DYRK1A is involved in BRCA1 chromatin recruitment, since the average BRCA1 ChIP read density is reduced upon DYRK1A depletion (Fig. R17). These results are compatible with either the need of DYRK1A on chromatin for the recruitment of BRCA1 to the co-occupied targets and/or with a DYRK1A-dependent effect that happens outside chromatin. Moreover, the results allow us to propose a model in which DYRK1A regulates BRCA1 nuclear localization by phosphorylation (Fig. D1). In fact, there are BRCA1-dependent processes that are regulated by phosphorylation besides those roles of BRCA1 in DDR and the control of cell cycle checkpoints. In one of them, phosphorylation of Thr509 by AKT/PKB enhances BRCA1 nuclear accumulation as well as its transcriptional activity (Altiok et al., 1999; Hinton et al., 2007). In the other, phosphorylation on Ser632 by CDK4 triggers BRCA1 chromatin eviction (Kehn et al., 2007). Our results show that DYRK1A phosphorylates BRCA1, at least, on Thr509 and Ser632 (Fig. R21). Therefore, it is possible that DYRK1A phosphorylation results in similar effects as the ones already described. However, whether these phosphorylation events take place *in situ* at chromatin has to be further studied.

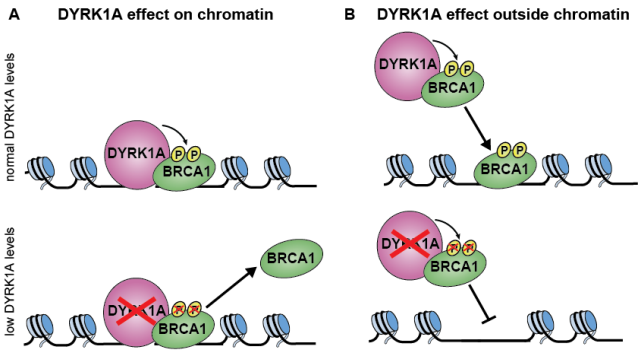


Figure D1: Schematic representation of the model for DYRK1A regulation of BRCA1 chromatin recruitment. (A-B) Proposed mechanism of DYRK1A regulation of BRCA1 chromatin recruitment directly on DYRK1A-BRCA1 co-occupied ChIP targets (**A**) or outside chromatin (**B**). See text for details.

One effect of the BRCA1-DYRK1A cross-talk on chromatin could be direct on the transcription of the common targets, based on their

enriched presence in promoter regions. However, BRCA1-DYRK1A co-occupancy at chromatin does not appear to have a major impact in the steady state levels of target transcripts (Fig. R22). We can envisage several explanations for these results. First, low BRCA1 levels on chromatin in DYRK1A or BRCA1 silenced conditions could be enough for BRCA1 transcriptional activity. Alternatively, it is possible that the nascent transcripts are the ones sensitive to acute BRCA1 chromatin eviction, a hypothesis that could be tested by NET-Seq or Global Run-on sequencing (GRO-Seq) analysis. In addition, it is possible that certain stimuli are required to trigger transcriptional changes. Since the main role attributed to BRCA1 is related to the DDR, DNA damage could be a proper scenario to approach. In fact, DYRK1A depletion has been related to increased radiation sensitivity and dysregulation of DSB-repair pathways (Roewenstrunk, 2016). Notably, some of the co-occupied target genes whose expression is affected by DYRK1A silencing are related to the DDR such as *POLE*, *RFC3* or *RFC5* (Fig. R23). In addition, BRCA1 has been shown to be recruited to R-loops formed at RNAPII pausing sites usually provoked by DNA damage (Hatchi et al., 2015); in this situation, ssDNA corresponding to the antisense strand is exposed for longer times, which could offer the possibility for DYRK1A recruitment to the palindromic motif.

In the context of BRCA1 chromatin recruitment and transcriptional activation, several issues should be mentioned when comparing our data with published results. First, our BRCA1 ChIP data show an enrichment in the DYRK1A-motif, which does not agree with a DNA sequence previously associated with BRCA1 binding, TTC(G/T)GTTG (Cable et al., 2003). Second, the BRCA1 ChIP datasets extracted from ENCODE show a higher number of BRCA1-associated regions than the ones obtained in this work (data not shown). Although this could be due to differences in the analysis procedure or to cell-type specific BRCA1 chromatin binding profiles, the possibility of being underestimating the number of BRCA1 ChIP-peaks should be contemplated as a plausible explanation for the poor overlap between BRCA1-related expression alterations and its chromatin occupancy. Third, two publications have analyzed BRCA1 recruitment to chromatin in a genome-wide manner (Gardini et al., 2014; Gorski et al., 2011). Similarly to our results, a low overlap between BRCA1 ChIP regions and BRCA1-dependent alterations in expression profiles of the corresponding gene targets was observed in both studies. The published data were generated from breast cancer cell lines and with two different antibodies than the one used in this Thesis work and in

ENCODE. Surprisingly, the overlap between the three BRCA1 binding profiles is very poor (data not shown). We have been able to compare the behavior of the antibody used by Gardini and colleagues (I20) with the one used in this Thesis work (000A) in a set of targets, and both correlate with the results obtained in the different ChIP-Seq datasets and to respond to BRCA1 silencing (Fig. D2). The results also provide further support to our data, since the targets that are positive for the antibody used in this work, are also positive for another BRCA1 antibody (D20) recognizing a different epitope. Although we cannot exclude the existence of technical issues, the results could suggest that different anti-BRCA1 antibodies might detect different pools of BRCA1 protein located at different genomic *loci*. Unfortunately, further studies using some of the antibodies will not be possible because they have been discontinued.

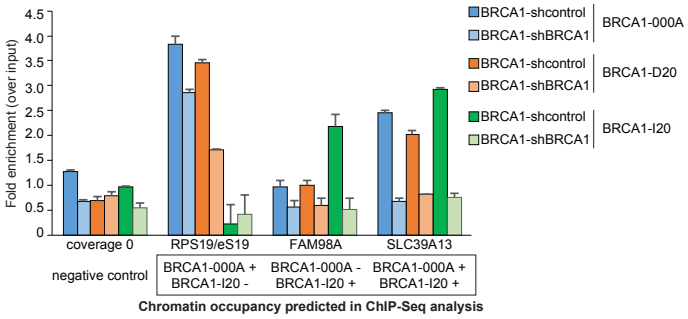


Figure D2: Validation of anti-BRCA1 antibodies for ChIP analysis. ChIP-qPCR experiment carried out in shControl *vs* shBRCA1 conditions using three independent anti-BRCA1 antibodies, BRCA1-000A (this Thesis work), BRCA1-D20 and BRCA1-I20 (Gardini et al., 2014) (see Fig. R5 for validation of all antibodies). Chromatin occupancy predictions extracted from ChIP-Seq data are indicated in the legend. A random genomic region (coverage 0) was selected as a negative control. The experiment was performed with an external spike-in as described in MM5.2 and 5.3. Data is represented as DNA recovery over the input (mean±SD of 3 technical replicates).

Although the effect of DYRK1A on BRCA1 chromatin recruitment is evident, the consequences on further molecular processes are still not clear and the different scenarios are just speculative. In addition to the already discussed effect on transcriptional regulation, one possibility is that both factors cooperate in co-transcriptional processes such as splicing. In this regard, BRCA1 has been demonstrated to interact with the splicing factor BCLAF1 in order to facilitate the mRNA processing of genes involved in the DDR (Savage et al., 2014). Interestingly, several DYRK1A substrates are related to splicing (Fig. I4). Therefore, it could be interesting to test whether BRCA1 and DYRK1A are acting

in co-transcriptional splicing in complexes containing BCLAF1 or other components of the splicing machinery.

Furthermore, DYRK1A-dependent BRCA1 chromatin ejection could have an impact in chromatin remodeling since both factors have been connected to this process via their interaction with the SWI/SNF complex (Harte et al., 2010; Lepagnol-Bestel et al., 2009). This hypothesis should be taken with caution since the data linking BRCA1 and DYRK1A with this complex were obtained from two different scenarios: estrogen receptor response in breast cells and dendritic growth in neurons, respectively and, in consequence, they might end up coordinating the regulation of different subsets of genes.

5. Chromatin-bound DYRK1A role on the transcriptional regulation of RPGs

5.1. Characterization of DYRK1A-positive RPG promoters

In the present work, we have demonstrated that DYRK1A is located at proximal promoters of a particular subset of RPGs and that it regulates their transcription likely at the chromatin level (Fig. R26 and R27). The DYRK1A-motif, TCTCGCGAGA, appears as a distinctive feature in the majority of DYRK1A-positive RPG promoters; it is therefore possible that DYRK1A occupies the promoters containing a consensus sequence in a cell-type independent manner, while occupancy of promoters containing degenerated versions of the consensus could be context-specific. We still do not know which nucleotides of the motif are important for the recruitment of the kinase. Therefore, we can only consider the conservation of the motif to establish different binding profiles, and further experiments are required to establish *bona fide* DYRK1A-consensus sites. On top of that, there are a few RPGs lacking the palindromic motif that have DYRK1A at their promoters, thus revealing a motif-independent DYRK1A chromatin recruitment mechanism, still to be elucidated.

Our data showed that BRCA1 and KAISO are present at motif-positive promoters of genes fulfilling functions related to ribosome processes (Fig. R9-R11), mostly represented by the DYRK1A-motif-positive RPG promoters (Fig. R28). To date, KAISO has not been described to be involved in translation-related processes, although we can establish certain links. Transcriptional regulation by KAISO contributes to oligodendrocyte differentiation (Zhao et al., 2016); in addition, it has been shown that during oligodendrocyte differentiation the transcription of a set of genes associated with ribosomal subunit biogenesis and assembly is induced (Pol et al., 2017). None of the publications provide

the corresponding lists of genes, but it is possible that the overlap exists. Regarding, the other physiological context in which KAISO has been previously associated with genomic *loci* enriched in the DYRK1A-motif, adipogenic differentiation, we have been able to confirm the presence of RPGs in the list of KAISO-associated targets; of note, translation has been recently shown to be regulated during the process (Reid et al., 2017), and RPG transcription has been observed to increase during the process (von der Heyde et al., 2014). We have not yet generated data of differentially expressed genes upon KAISO silencing, so a proper evaluation of the association has to be done in the future.

In the case of BRCA1, and despite BRCA1 presence at the subset of RPG promoters, we could not observe any changes in the mRNA expression of the corresponding RPGs upon BRCA1 knock-down. However, there are some evidence pointing to a role of BRCA1 in ribosome biogenesis. On one side, it has been shown that BRCA1 is recruited to rDNA repeats during DNA damage and regulates rRNA transcription (Johnston et al., 2016); on the other, BRCA1 has been shown to regulate translation in breast cancer cells (Dacheux et al., 2013). Therefore, it could be interesting to test whether an effect in RPG expression is observed upon DNA damage as a regulatory process of ribosome functionality.

Interestingly, 60% of motif-positive RPG promoters contain a binding site for NR4A1, an orphan nuclear receptor (Fig. R30A). This TF is mainly involved in the regulation of genes related to inflammatory responses, oxidative metabolic pathways and apoptosis, thereby having an impact in cell proliferation and differentiation (Beard et al., 2015). Current data indicates that NR4A1 might play opposite roles in cell survival and tumorigenesis, either as an activator of pro-apoptotic pathways (Deutsch et al., 2014; Wu et al., 2017), or as a pro-oncogenic factor (Hedrick et al., 2017; Lee et al., 2017). In addition, NR4A1 has been observed to induce protein synthesis via activating mTOR pathway either indirectly via p53 inhibition or directly, by forming a trimeric complex with the components of the cascade Tuberous sclerosis protein 1/2 (Lee et al., 2012; Wang et al., 2013). Despite the enrichment in NR4A1 binding sites, analysis of published data (Duren et al., 2016) showed low levels of occupancy at RPG promoters in the Kasumi cell line (Fig. R30C). This inconsistency could be due to a cell-type specific presence of NR4A1 at genomic regions, and therefore, it could be interesting to investigate the NR4A1 chromatin profile in T98G. If true, NR4A1 and DYRK1A may play co-regulatory roles in RPG expression.

Our analysis also indicates that ETS binding sites are highly present in the cluster of RPGs lacking the DYRK1A-motif (Fig. R30B). In agreement, GABP, an ETS family member, occupies proximal promoter regions of RPG promoters depleted of DYRK1A and its occupancy is low in DYRK1A-positive RPG promoters (Fig. R30C), suggesting that these two proteins might mark different subsets of RPGs to be differentially regulated. In this context, changes in the transcripts of 14 RPs have been detected in the brain of DS mouse models (Chrast et al., 2000). Like *DYRK1A*, *GABP* is located in human chromosome 21. Therefore, while the imbalance in *DYRK1A* gene dosage may explain transcriptional dysregulation in those RPGs containing the kinase at their promoters, an extra copy of *GABP* could explain the alterations found in the expression of DYRK1A-negative RPGs.

In conclusion, the new information that we have generated on the characterization of RPG promoters may help understanding how RPG expression is regulated.

5.2. DYRK1A-dependent differential transcription of RPGs may affect the function of specialized ribosomes and/or ribosome concentration

As indicated in the Introduction (I3.3), a novel level of gene expression regulation is associated with the differential translation of mRNAs as a consequence of variations found in the composition of ribosomes. In the present work, we have demonstrated that a particular subset of RPGs, having DYRK1A at their promoters, are downregulated upon DYRK1A depletion (Fig. R27), and we propose that this process could, at least in part, contribute to the global reduction in protein synthesis observed when DYRK1A is silenced (Fig. R24). Currently, we cannot distinguish whether this particular process causes a global defect or just affects a subset of mRNAs. As part of a combinatorial process contributing to the final phenotype, it is possible that the DYRK1A-dependent effect on RP stoichiometry and the subsequent differential translation of mRNAs may be masked in the observed global reduction in translational rate. In this line, we have observed that some protein bands appeared to be differentially affected, in the ^{35}S-Met incorporation assay following DYRK1A reduction of expression (Fig. R24B).

In any case, differences in translatability, as measured in the assays shown in Fig. R24, could be due to several mechanisms acting independently or in concert: i) different transcript levels, ii) differential

recruitment of certain mRNAs to polysomes, iii) changes in ribosome concentration, since poorly translated mRNAs could be more sensitive to a reduction in ribosome concentration (Mills and Green, 2017), and iv) the existence of specialized ribosomes. For instance, and in the context of specialized ribosomes, it has been described that downregulation of a DYRK1A-positive RPG, RPL10A/uL1, leads to a decrease in the translatability of a specific pool of mRNAs characterized by the presence of IRES elements in their 5'-UTRs and that the specific recognition of these mRNAs was explained by RPL10A/uL1 location at the mRNA exit tunnel (Shi et al., 2017). In this context, RPs containing DYRK1A at their promoters are equally found in the two ribosomal subunits and homogenously distributed within the ribosome structure (Fig. D3) so that extrapolating structural functions of this particular set of RPs as a group is not straightforward.

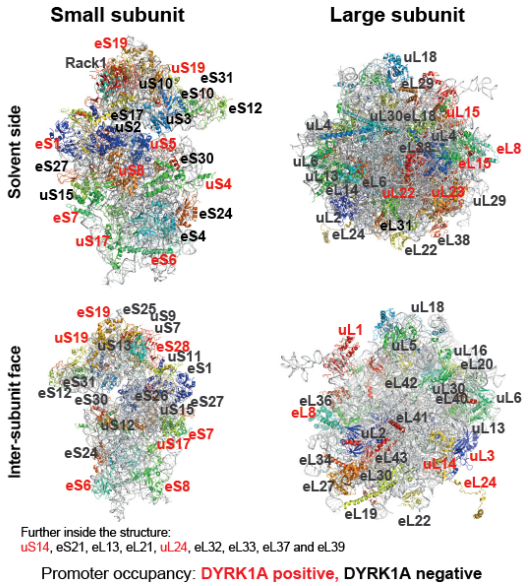


Figure D3: Distribution of RPs within the ribosome structure. Global view of both subunits conforming the 80S human ribosome (small subunit: left; large subunit: right) from the solvent side (upper part) and from the inter-subunit interface (lower part) with RPs indicated on the atomic model. Proteins uS14, eS21, eL13, eL21, uL24, eL32, eL33, eL37 and eL39 are further inside the structure and are not labeled. Those proteins whose promoters present DYRK1A are highlighted in red. RPs are named according to the new nomenclature proposed in Ban et al., 2014 (equivalences to the old nomenclature are listed in Annex I). Adapted from Khatter et al., 2015.

In summary, we propose that a putative DYRK1A-dependent reduction either in specific RPs or in global ribosome content would lead to a differential impact in mRNA translatability. To provide some answers, the pool of mRNAs associated with polysomes that respond

differentially to DYRK1A downregulation needs to be identified; this information together with the identification of *cis*-regulatory elements in these transcripts will surely help to discriminate among several of the possibilities and provide a mechanistic framework.

5.3. DYRK1A-dependent differential transcription of RPGs may affect extra-ribosomal functions

Several RPs exert functions out of the ribosome; therefore, DYRK1A impact on RP levels may be connected to some of these functions. As RNA binding proteins, RPs are involved in rRNA processing (Xue and Barna, 2012) and this would be in agreement with rRNA binding process as one of the most enriched categories appearing in the GO terms analysis of DYRK1A ChIP targets downregulated in shDYRK1A conditions (Fig. R23). Therefore, DYRK1A may alter ribosome functionality by playing a role at other regulatory layers.

Apart from that, p53-dependent cell cycle arrest is considered as the most known extra-ribosomal function. Free RPL5/uL18 and RPL11/uL5 are the main effectors triggering p53 activation via MDM2/HDM2 binding (Bursac et al., 2012). Although neither of these RPs are direct DYRK1A ChIP targets, defects in the expression of some RPs as a consequence of DYRK1A absence could promote an imbalance in RP incorporation into ribosomes and an increase in free RPs. DYRK1A depletion has been previously demonstrated to lead to p53 activation by phosphorylation, and therefore to cell cycle arrest (Park et al., 2010). In fact, this could be a contributing factor to the G1 increase showed by DYRK1A-silenced cells (Fig. R24A). Taking into account these results, a model can be proposed in which DYRK1A prevents p53 activation by preserving the integrity of ribosomes (Fig. D4).

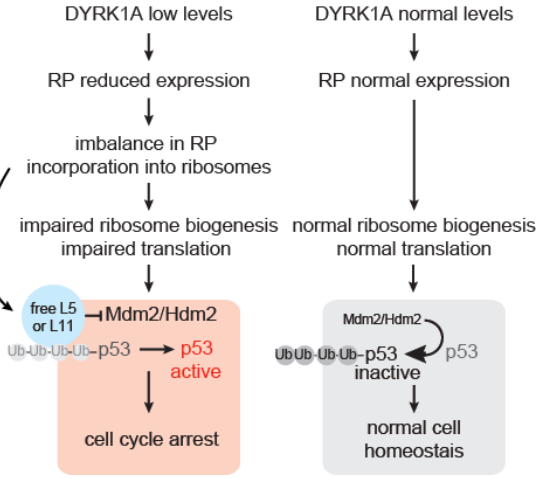


Figure D4: Model for DYRK1A acting on p53 activation by preserving the integrity of ribosomes. See text for further details.

6. Physiological links between DYRK1A and ribosome-defective phenotypes

We propose that the defect in protein synthesis upon DYRK1A downregulation is partially due to a reduction in the transcription of a subset of RPs in a DYRK1A-dependent manner. An interesting question is how extended is the phenomenon and whether this process has an impact in physiological contexts. Regarding the first question, the expression of 5 DYRK1A-positive RPGs has been found altered in the brains of DS mouse models (Chrast et al., 2000).

With respect to the second question, an immediate consequence of defects in RP expression and protein translation would be a reduction in cell mass growth, which is in concordance with the fact that DYRK1A-depleted T98G cells present a reduction in cell volume (Di Vona, 2013). Size control is a complex process in higher eukaryotic cells, in which both intrinsic and extrinsic factors appear to contribute ([Amodeo and Skotheim, 2016; Ginzberg et al., 2015).

Based on current knowledge, we envisage two scenarios in which DYRK1A control over cell size could be important. One of them is related to neuron differentiation, since huge postmitotic growth efforts are required for the neuron terminals to reach their targets far away from their position. It is well known that DYRK1A plays essential roles in central nervous system development, not only affecting cell numbers but also differentiation processes (Guedj et al., 2012; Martinez de Lagran et al., 2012). Therefore, the effect of DYRK1A on translation might contribute to its role as a regulator of neurite and axonal growth. In another context, it is known that in pancreatic β–cells, insulin secretion, metabolic activity, and global rates of protein production are all correlated with cell size (Ruvinsky et al., 2005). Although the effect of DYRK1A on β–cell homeostasis is still controversial, mouse models overexpressing *Dyrk1a* show expansion of β–cell mass through increased proliferation and cell size (Rachdi et al., 2014). It is possible that the DYRK1A impact in cell mass may also apply to additional tissues since heterozygous mouse models exhibit a global reduction in body size (Fotaki et al., 2002). Interestingly, all the organs in the heterozygous mice were smaller compared to control animals with the exception of the kidney, where DYRK1A was very low expressed (unpublished results; see Fig. I3 for DYRK1A mRNA levels in human tissues). These observations point to a role of DYRK1A as a regulator of cell growth and suggest that the DYRK1A-dependent regulation of RP expression could be one of the processes contributing to the final phenotype.

Furthermore, defects in protein production are closely related to cancer since increased translation is required to boost cell proliferation (further detailed in I3.4). In fact, RPs are considered oncogenic drivers (Sulima and De Keersmaecker, 2017). The role of DYRK1A in cancer is controversial, since it can negatively or positively affect cell proliferation depending on the tumor context (reviewed in I1.3), but DYRK1A-dependent transcriptional activation of RPs could be considered as a new mechanism contributing to cancer development.

Finally, mutations in specific RPGs can lead to very unique phenotypes. *RPL5/uL18* represents an interesting example since, although its promoter does not contain DYRK1A, it is a particular case found to be significantly downregulated in DYRK1A silenced conditions. Mutations in *RPL5/uL18* lead to craniofacial anomalies, hearth alterations and tumor predisposition and the two first features are hallmarks of *DYRK1A* haploinsufficiency syndrome or of animal models with *Dyrk1a* dysregulation. In conclusion, DYRK1A might play a role in specific ribosome-related alterations via affecting RPG expression.

7. Final remarks: DYRK1A in the frame of cell growth

Finally, I would like to finish the Discussion section highlighting the DYRK1A role in cell growth regulation, and I would like to propose that it is very likely exerted at several regulatory layers, some of them already known and some others to be discovered in case proteins involved in translational control resulted to be DYRK1A substrates.

Although DYRK1A participation in the RP expression regulation may play an important role in the process, DYRK1A might exert additional functions related to the control of translation. At the chromatin level, DYRK1A has been demonstrated to regulate the transcription of RNA binding proteins involved in the maturation of rRNA (NOL11, UTP6) (Fig. R23), as well as some translation factors (DENR) (Di Vona, 2013). In addition, work performed in our laboratory showed that DYRK1A is recruited to regions encoding tRNAs, essential tools in the translation process. Moreover, they seem to have a DYRK1A-dependent transcriptional activation (Di Vona, 2013). Therefore, chromatin-bound DYRK1A may affect protein synthesis by regulating the transcription of different components essential for ribosome biogenesis and functionality (Fig. D5).

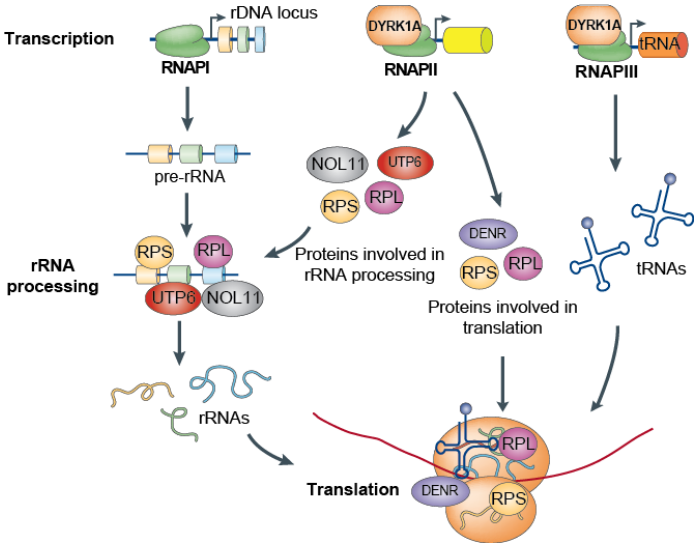


Figure D5: Schematic representation of chromatin-bound DYRK1A functions playing a regulatory role in translation. See text for further details.

Interestingly, not only the activation of determinant signaling pathways and transcriptional programs controls protein synthesis but, in turn, the translational status may affect chromatin dynamics within a cell, specially in situations of hypertranscription such as in the frame of ESCs (Bulut-Karslioglu et al., 2018). Since DYRK1A seems to be an important regulator of cell growth and protein synthesis, it could be interesting to assess whether DYRK1A downregulation is affecting the epigenetic landscape as a consequence of fluctuations in the translational rate of the cells. To date, two chromatin marks, phosphorylation at Thr45 and acetylation of H3, have been analyzed at several DYRK1A-associated promoters and no changes were observed upon DYRK1A depletion in T98G cells (Di Vona et al., 2015). This does not imply that DYRK1A levels might modulate some other histone marks at genome-wide levels and in a transcriptionally permissive state such as in the case of ESCs, a possibility that deserves to be studied in the future.

Furthermore, there is a close link between the regulatory mechanisms controlling cell size and cell number. On one hand, the cell must reach a minimal size to undergo mitosis and, on the other hand, the first consequence of cell division is a reduction in cell volume (Schmoller, 2017). In fact, it has been observed negative correlation between cell size and G1 length (Liu et al., 2018). In this regard, it is going to be difficult to untangle whether the reduced cell volume observed in DYRK1A-depleted cells is what makes the cells unable to bypass the

G1/S checkpoint or if the impact of DYRK1A on cell cycle regulators such as Cyclin D1 or p27 induces the G1 arrest and subsequently the reduction in cell size. Noteworthy, targeting DYRK1A with two independent shRNAs leads to a phenotype defective in RPG expression and protein synthesis (Fig. R24C and R27B). By contrast, the cell cycle profile is altered when the expression of the kinase is reduced with shDYRK1A.1 (Fig. R24A), but no alterations were described when using shDYRK1A.2 (Di Vona, 2013). These results suggest that the defects observed in translation and cell volume upon DYRK1A silencing are independent of the defects in cell cycle.

Ultimately, the mechanisms by which a cell senses its own size and is able to adapt its physiology according to different circumstances are not well understood, and there are still many questions to answer in the field. In this Thesis work, we describe DYRK1A as a novel transcriptional activator of RP expression. Evidence is accumulating on DYRK1A contributing to cell growth control at different regulatory steps, a situation that resembles in many aspects the activity of another kinase, mTOR, as a master regulator of this process (Giguere, 2018; Saxton and Sabatini, 2017). Therefore, future investigations will help to dissect DYRK1A activities in the nucleus and cytoplasm, how they are linked and how they converge in controlling cell growth.

Conclusions

1. DYRK1A interacts directly with single strand oligonucleotides containing the TCTCGCGAGA palindromic sequence *in vitro.*

2. The chromatin binding profile obtained with the anti-CHD2 antibody ab68301 used by the ENCODE Consortium overlaps with that of DYRK1A in T98G cells. However, the results proved that this antibody does not recognize CHD2 but an unknown protein.

3. KAISO and BRCA1 chromatin occupancies mostly overlap with DYRK1A-occupied genomic regions in T98G cells.

4. DYRK1A, KAISO and BRCA1 co-occupied genomic regions are enriched in TCTCGCGAGA-containing promoters associated to ribosome-related functions.

5. There is not a clear interdependency between DYRK1A and KAISO for chromatin recruitment.

6. BRCA1 chromatin recruitment is negatively affected by the reduction of DYRK1A expression.

7. DYRK1A and BRCA1 are part of common complexes in the soluble nuclear fraction.

8. BRCA1 is a DYRK1A substrate, with several DYRK1A-dependent phosphosites scattered along BRCA1 primary sequence.

9. The cross-talk between BRCA1 and DYRK1A at the promoter of common target genes does not result in changes in the steady-state levels of the corresponding transcripts in normal growing conditions.

10. DYRK1A depletion leads to a reduction in the translational rate without affecting eIF2α-dependent pathways.

11. DYRK1A, BRCA1 and KAISO are present at a cluster of ribosomal protein gene promoters enriched in a highly conserved TCTCGCGAGA palindromic sequence.

12. The presence of DYRK1A at the promoters of ribosomal protein genes positively correlates with their transcription levels.

13. Ribosomal protein gene promoters containing the DYRK1A-consensus are enriched in NR4A1 binding sites.

14. GABP is differentially located at ribosomal protein gene promoters depleted of DYRK1A, suggesting the existence of two clusters of RPGs from the viewpoint of their transcriptional regulation.

Abbreviations

aa: Amino acid/s
ABC: Ammonium bicarbonate
Abf1: ARS-binding factor 1
Ablim: Actin binding LIM protein
AGC: Autogain control
AMP: Adenosine monophosphate
AP4: Activator protein 4
App: Amyloid precursor protein
Arip4: Androgen receptor interacting protein 4
ASD: Autism spectrum disorders
ATM: Ataxia-telangiectasia mutated
ATP: Adenosine triphosphate
ATR: Ataxia-telangiectasia Rad3-related
BARD1: BRCA1-associated RING domain protein 1
BCLAF1: B-cell lymphoma 2-associated factor 1
Bp: Base pair
BRCA1: Breast cancer type 1 susceptibility protein
BRCA2: Breast cancer type 2 susceptibility protein
BRCT: BRCA1 C-terminal
BRD4: Bromodomain containing protein 4
BRE: TFIIB recognition element
BS: Binding site
BSA: Bovine Serum Albumin
BTB/POZ: Broad complex, Trantraj, Bruc à brac/ Pox virus and zinc finger
CAK: Cdk-activating kinase
CDC37: Cell division cycle 37.
CDK: Cyclin-dependent kinase
CDKL: Cyclin-dependent like kinase
CEAS: Cis-regulatory elements annotation system
CHD2: Chromodomain helicase DNA binding protein 2
ChIP: Chromatin Immunoprecipitation
Chk2: Checkpoint kinase 2
CHX: Cycloheximide
CK2: Casein kinase 2
CLK: cdc2-like kinase
CMV: Cytomegalovirus
Cp: Crossing point
CpG: Cytosine/Guanine dinucleotide
CPM: Counts per million
CRE: cAMP responsive element
CREB1: cAMP responsive element binding 1
Crf1: Co-repressor with FHL1
CRY2: Cryptochrome circadian clock 2
CTCF: CCCTC-binding factor
CTD: Carboxy-terminal domain
C-terminal: Carboxy-terminal
Cys: Cysteine
DAPI: 4’,6-diamino-2-phenylindole
DBD: DNA binding domain

DCAF7: DNA damage-binding protein and cullin 4-associated protein 7
DCE: Downstream core element
DDA: Data dependent Acquisition
DDR: DNA damage response
DENR: Density-regulated protein
DEXDc: Dead-like helicase domain
DH: DYRK homology
DMEM: Dulbecco's modified Eagle's medium
DNA: Desoxiribonucleic acid
dNTP: Deoxyribonucleotide
DPE: Downstream promoter element
DRE: DNA replication element
DREAM: Dimerization partner, RB-like, E2F and multi-vulval class B
DS: Down syndrome
DSB: Double strand break
DSCR: Down syndrome critical region
DSIF: DRB sensitivity-inducing factor
DTT: Dithiothreitol
DYRK: Dual-specificity tyrosine(Y)-regulated kinase
DH-box: DYRK-homology box
EDTA: Ethylendiamine tetracetic acid
EEF1A1: Eukaryotic elongation factor 1 alpha 1
EF: Elongation factor
EGCG: Epigallocatechin gallate
EGFR: Epidermal growth factor receptor
eIF: Eukaryotic initiation factor
EMSA: Electrophoretic mobility shift assay
ENAH/MENA: Enabled homolog/Mammalian enabled
ENCODE: Encyclopedia of DNA elements
ER: Estrogen receptor
eRF: Eukaryotic release factor
ERK: Extracellular signal-regulated kinase
ESC: Embryonic stem cell
ETD: Electron transfer dissociation
ETS: E26-transformation specific/ E-Twenty-Six
FAM98A: Family with sequence similarity 98 member A
FBS: Fetal bovine serum
FC: Fold change
FDR: False discovery rate
Fhl1: Forkhead-like 1
FOXO1: Forkhead box O1
FPKMs: Fragments per kilobase of transcript per million mapped reads
FRK: Fyn-related kinase
GABP: GA-binding protein
GCN2: General control nonderepressible 2
GEO: Gene expression omnibus
GFP: Green fluorescent protein
GLI1: Glioma-associated oncogene
GO: Gene ontology

GRO-Seq: Global run-on sequencing
GSK: glycogen synthase kinase
GST: Glutathione-S-transferase
GTex: Genotype-tissue expression
GTF-II: General transcription factor – II
GTP: Guanosine-5'-triphosphate
H3: Histone 3
HCD: High-energy collision dissociation
HCT: Human colorectal carcinoma
HDAC: Histone deacetylase
hDREF: Human DNA replication element factor
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HeLa: Henrietta Lacks
Hip1: Huntingtin interacting protein 1
HIPK: Homeodomain-interacting kinase
His: Histidine
HMO1: High mobility group protein 1
HNE: HeLa nuclear extract
HOX: Homeotic, homeobox
HP1: Heterochromatin protein 1
HRP: Horseradish peroxidase
HSP90: Heat shock protein 90
IC: Initiation complex
ICRI: Imprinting control region I
ID2: Inhibitor of DNA binding 2
Ifh1: Interacts with Forkhead 1
IgGs: Immunoglobulins G
Inr: Initiator
IP: Immunoprecipitation
IPTG: Isopropyl-β-D-1-thiogalactopyranoside
IQR: Interquartile range
IRES: Internal ribosome entry site
IVK: *In vitro* kinase assay
JNK: c-Jun amino (N)-terminal kinase
Kb: Kilobase
kDa: kilodalton
LATS2: Large tumor suppressor kinase 2
LB: Lysogeny broth
LiCl: Lithium Chloride
LncRNA: Long non-coding RNA
LoF: Loss-of-function
MAP1B: Microtubule-associated protein 1B
MAPK: Mitogen-activated protein kinase
mCpG: methylated Cytosine/Guanine dinucleotide
MDC1: Mediator of DNA damage checkpoint 1
MDM2/HDM2: Murine/human double minute 2
MEF2D: Myocyte-specific enhancer factor 2
Met: Methionine
Min: Minute/s

Mnb: Minibrain
mIgG: mouse immunoglobulin G
miRNA: Micro-RNA
mRNA: Messenger RNA
MS: Mass spectrometry, mass spectrometer
MTA2: Metastasis associated protein 2
mTOR: Mammalian target of rapamycin
Munc18-1: Syntaxin binding protein
MUT: Mutated
NAPA: N-terminal autophosphorylation accessory region
NCBI: National Centre for Biotechnology Information
N-CoR: Nuclear receptor co-repressor 1
NELF: Negative elongation factor
NET-Seq: Native elongating transcript sequencing
NFAT: Nuclear factor of activated T-cells
NLK: Nemo-like kinase
NLS: Nuclear localization signal
NMDA2A: Glutamate ionotropic receptor NMDA type subunit 2A
NP-40: nonidet P-40
NOL11: Nucleolar protein 11
NR4A1: Nuclear receptor subfamily 4 group A member 1
N-terminal: Amino-terminal
N-WASP: Neural Wiskott-Aldrich syndrome protein
OMIM: Online mendelian inheritance in man
O/N: Overnight
ORPHA: Orphanet
PAGE: Polyacrylamide gel electrophoresis
PAHX-AP1: Phytanoyl-CoA alpha-hydroxylase-associated protein 1
PALB2: Partner and localizer of BRCA2
PAS: PolyA site
PBS: Phosphate-buffered saline
PCR: Polymerase chain reaction
PERK: Protein kinase RNA-like endoplasmic reticulum kinase
PEST: Region rich in proline, glutamic acid, serine and threonine residues
Phe: Phenylalanine
PIC: Pre-initiation complex
PKB/AKT: Protein kinase B
PKR: Protein kinase double-stranded RNA-dependent
PLK: Polo-like kinase
POLE: DNA polymerase epsilon catalytic subunit
POLR2: RNA polymerase II subunit A
PPIA: Peptidyl-prolyl-isomerase A
Pro: Proline
PRP4K: Pre-mRNA processing protein 4 kinase
pTEFb: Positive transcription elongation factor b
qPCR: Quantitative polymerase chain reaction
RAD51: Recombinase RAD52, RecA homolog
Rap1: Repressor activator protein 1
Rb: Retinoblastoma

RBM: RNA binding motif protein
RCAN1: Regulator of calcineurin 1
REST: RE1 silencing transcription factor
RFC: Replication factor C
RIN: RNA integrity number
RING: Really interesting new gene
RNA: Ribonucleic acid
RNF: Ring finger
RNAPI: RNA polymerase I
RNAPII: RNA polymerase II
RNAPIII: RNA polymerase III
RP: Ribosomal protein
RPB1: RNA-directed RNA polymerase II subunit 1
RPG: Ribosomal protein gene
rRNA: Ribosomal RNA
RT-qPCR: Real Time-quantitative Polymerase Chain Reaction
SD: Standard deviation
SDS: Sodium dodecil sulfate
Seq: Sequencing
Ser (S): serine
SF3B1: Splicing factor 3b subunit 1
Sfp1: Split finger protein 1
shRNA: Short hairpin RNA
siRNA: Small interfering RNA
SIRT1: Sirtuin 1
SLC39A13: Solute carrier family 39 member 13
SMAD1: Mothers against decapentaplegic homolog 1
SMRT: silencing mediator of retinoic acid and thyroid hormone receptor
SNF2: Sucrose non-fermentable 2
snRNA: Small nuclear RNA
snoRNA: Small nucleolar RNA
snRNP: Small nuclear ribonucleoprotein particles
SP: Specificity protein
SPRED: Sprouty-related, EVH1 domain-containing protein
SR: Serine/arginine rich protein
SRSF: Serine-arginine-rich splicing factor
SRPK: Serine-arginine-rich protein kinase
ssDNA: Single strand DNA
SWI/SNF: Switch/Sucrose non-fermentable
TAF: TBP-associated factor
TBE: Tris/Borate/EDTA
TBP: TATA-Binding Protein
TBPL1: TBP-Like 1
TBX5: T-box transcription factor 5
TCF: T-cell factor
TE: Tris-ethylendiamine tetracetic acid
TF: transcription factor
Thr (T): Threonine
TOP: Track of polypyrimidine

TORC1: Target of Rapamycin complex 1
TPM: Transcripts per million
TRF2: TBP-related factor 2
TSS: Transcription start site
tRNA: Transfer RNA
Tyr: Tyrosine
Ubc13: Ubiquitin-conjugating enzyme E2 13
UCSC: University of California Santa Cruz
UTP: U3 small nucleolar RNA-associated protein
UTR: Untranslated region
VASP: Vasodilator-stimulated phosphoprotein
WB: Western blot
Wig: Wiggle
WT: Wild-type
Xrn2: Exoribonuclease 2
YY1: Yin Yang 1
ZBED1: Zinc finger BED-type containing 1
ZBTB33: Zinc finger and BTB containing domain protein 33
ZF: Zing finger

References

Adayev T, Chen-Hwang MC, Murakami N, Lee E*, et al.* 2007. Dual-specificity tyrosine phosphorylation-regulated kinase 1A does not require tyrosine phosphorylation for activity in vitro. Biochemistry *46*, 7614-7624.

Albert B, Knight B, Merwin J, Martin V*, et al.* 2016. A molecular titration system coordinates ribosomal protein gene transcription with ribosomal RNA synthesis. Mol Cell *64*, 720-733.

Altafaj X, Dierssen M, Baamonde C, Marti E, Visa J*, et al.* 2001. Neurodevelopmental delay, motor abnormalities and cognitive deficits in transgenic mice overexpressing Dyrk1A minibrain, a murine model of Down's syndrome. Hum Mol Genet *10*, 1915-1923.

Altiok S, Batt D, Altiok N, Papautsky A*, et al.* 1999. Heregulin induces phosphorylation of BRCA1 through phosphatidylinositol 3-kinase/AKT in breast cancer cells. J Biol Chem *274*, 32274-32278.

Alvarez M. 2004. Localización subcelular de la proteína quinasa DYRK1A: compartimentos, señales y regulación. Universitat de Barcelona, Spain.

Alvarez M, Altafaj X, Aranda S, and de la Luna S. 2007. DYRK1A autophosphorylation on serine residue 520 modulates its kinase activity via 14-3-3 binding. Mol Biol Cell *18*, 1167-1178.

Alvarez M, Estivill X, and de la Luna S. 2003. DYRK1A accumulates in splicing speckles through a novel targeting signal and induces speckle disassembly. J Cell Sci *116*, 3099-3107.

Amodeo AA, and Skotheim JM. 2016. Cell-Size Control. Cold Spring Harb Perspect Biol *8*, a019083.

Andreev DE, O'Connor PB, Loughran G, Dmitriev SE*, et al.* 2017. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. Nucleic Acids Res *45*, 513-526.

Aranda S, Alvarez M, Turro S, Laguna A, and de la Luna S. 2008. Sprouty2-mediated inhibition of fibroblast growth factor signaling is modulated by the protein kinase DYRK1A. Mol Cell Biol *28*, 5899-5911.

Aranda S, Laguna A, and de la Luna S. 2011. DYRK family of protein kinases: evolutionary relationships, biochemical properties, and functional roles. FASEB J *25*, 449-462.

Arato K. 2010. Regulation of the stability of the protein kinase DYRK1A: establishing connections with the Wnt signaling pathway. Universitat Pompeu Fabra, Barcelona, Spain.

Arranz J. 2016. Caracterización de las alteraciones estructurales y funcionales de la corteza cerebral de ratones mutantes de pérdida y ganancia de función de DYRK1A. Universidad de Barcelona, Spain.

Arron JR, Winslow MM, Polleri A, Chang CP*, et al.* 2006. NFAT dysregulation by increased dosage of DSCR1 and DYRK1A on chromosome 21. Nature *441*, 595-600.

Bailey TL, Boden M, Buske FA, Frith M*, et al.* 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Bain J, Plater L, Elliott M, Shpiro N*, et al.* 2007. The selectivity of protein kinase inhibitors: a further update. Biochem J *408*, 297-315.

Ban N, Beckmann R, Cate JH, Dinman JD*, et al.* 2014. A new system for naming ribosomal proteins. Curr Opin Struct Biol *24*, 165-169.

Barallobre MJ, Perier C, Bove J, Laguna A*, et al.* 2014. DYRK1A promotes dopaminergic neuron survival in the developing brain and in a mouse model of Parkinson's disease. Cell Death Dis *5*, e1289.

Bartkowiak B, Liu P, Phatnani HP, Fuda NJ*, et al.* 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. Genes Dev *24*, 2303-2316.

Baumann DG, and Gilmour DS. 2017. A sequence-specific core promoter-binding transcription factor recruits TRF2 to coordinately transcribe ribosomal protein genes. Nucleic Acids Res *45*, 10481-10491.

Beard JA, Tenga A, and Chen T. 2015. The interplay of NR4A receptors and the oncogene-tumor suppressor networks in cancer. Cell Signal *27*, 257-266.

Becker W, and Joost HG. 1999. Structural and functional characteristics of Dyrk, a novel subfamily of protein kinases with dual specificity. Prog Nucleic Acid Res Mol Biol *62*, 1-17.

Becker W, and Sippl W. 2011. Activation, regulation, and inhibition of DYRK1A. FEBS J *278*, 246-256.

Becker W, Weber Y, Wetzel K, Eirmbter K*, et al.* 1998. Sequence characteristics, subcellular localization, and substrate specificity of DYRK-related kinases, a novel family of dual specificity protein kinases. J Biol Chem *273*, 25893-25902.

Beishline K, and Azizkhan-Clifford J. 2015. Sp1 and the 'hallmarks of cancer'. FEBS J *282*, 224-258.

Belgardt BF, and Lammert E. 2016. DYRK1A: A promising drug target for islet transplant-based Diabetes therapies. Diabetes *65*, 1496-1498.

Benjamini YHY. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc *57*, 289-300.

Bhat M, Robichaud N, Hulea L, Sonenberg N*, et al.* 2015. Targeting the translation machinery in cancer. Nat Rev Drug Discov *14*, 261-278.

Blattler A, Yao L, Wang Y, Ye Z*, et al.* 2013. ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes. Epigenetics Chromatin *6*, 13.

Blazek JD, Abeysekera I, Li J, and Roper RJ. 2015. Rescue of the abnormal skeletal phenotype in Ts65Dn Down syndrome mice using genetic and therapeutic modulation of trisomic Dyrk1a. Hum Mol Genet *24*, 5687-5696.

Bochar DA, Wang L, Beniya H, Kinev A*, et al.* 2000. BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. Cell *102*, 257-265.

Bohne F, Langer D, Martine U, Eider CS*, et al.* 2016. Kaiso mediates human ICR1 methylation maintenance and H19 transcriptional fine regulation. Clin Epigenetics *8*, 47.

Bonhoure N, Bounova G, Bernasconi D, Praz V*, et al.* 2014. Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization. Genome Res *24*, 1157-1168.

Bornelov S, Komorowski J, and Wadelius C. 2015. Different distribution of histone modifications in genes with unidirectional and bidirectional transcription and a role of CTCF and cohesin in directing transcription. BMC Genomics *16*, 300.

Bosio MC, Fermi B, and Dieci G. 2017. Transcriptional control of yeast ribosome biogenesis: A multifaceted role for general regulatory factors. Transcription *8*, 254-260.

Branca C, Shaw DM, Belfiore R, Gokhale V*, et al.* 2017. Dyrk1 inhibition improves Alzheimer's disease-like pathology. Aging Cell *16*, 1146-1154.

Bulut-Karslioglu A, Macrae TA, Oses-Prieto JA, Covarrubias S*, et al.* 2018. The transcriptionally permissive chromatin state of embryonic stem cells is acutely tuned to translational output. Cell Stem Cell *22*, 369-383 e368.

Bursac S, Brdovcak MC, Pfannkuchen M, Orsolic I*, et al.* 2012. Mutual protection of ribosomal proteins L5 and L11 from degradation is essential for p53 activation upon ribosomal biogenesis stress. Proc Natl Acad Sci USA *109*, 20467-20472.

Bustelo XR, and Dosil M. 2018. Ribosome biogenesis and cancer: basic and translational challenges. Curr Opin Genet Dev *48*, 22-29.

Cable PL, Wilson CA, Calzone FJ, Rauscher FJ, 3rd*, et al.* 2003. Novel consensus DNA-binding sequence for BRCA1 protein complexes. Mol Carcinog *38*, 85-96.

Cen L, Xiao Y, Wei L, Mo M*, et al.* 2016. Association of DYRK1A polymorphisms with sporadic Parkinson's disease in Chinese Han population. Neurosci Lett *632*, 39-43.

Chabalier-Taste C, Brichese L, Racca C, Canitrot Y*, et al.* 2016. Polo-like kinase 1 mediates BRCA1 phosphorylation and recruitment at DNA double-strand breaks. Oncotarget *7*, 2269-2283.

Chen EY, Tan CM, Kou Y, Duan Q*, et al.* 2013a. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Chen JY, Lin JR, Tsai FC, and Meyer T. 2013b. Dosage of Dyrk1a shifts cells within a p21-cyclin D1 signaling map to control the decision to enter the cell cycle. Mol Cell *52*, 87-100.

Chen K, Hu Z, Xia Z, Zhao D*, et al.* 2015. The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses. Mol Cell Biol *36*, 662-667.

Chenier S, Yoon G, Argiropoulos B, Lauzon J*, et al.* 2014. CHD2 haploinsufficiency is associated with developmental delay, intellectual disability, epilepsy and neurobehavioural problems. J Neurodev Disord *6*, 9.

Choudhury AD, Xu H, and Baer R. 2004. Ubiquitination and proteasomal degradation of the BRCA1 tumor suppressor is regulated during cell cycle progression. J Biol Chem *279*, 33909-33918.

Chrast R, Scott HS, Papasavvas MP, Rossier C*, et al.* 2000. The mouse brain transcriptome by SAGE: differences in gene expression between P30 brains of the partial trisomy 16 mouse model of Down syndrome Ts65Dn and normals. Genome Res *10*, 2006-2021.

Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Cortez D, Wang Y, Qin J, and Elledge SJ. 1999. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. Science *286*, 1162-1166.

da Costa Martins PA, Salic K, Gladka MM, Armand AS*, et al.* 2010. MicroRNA-199b targets the nuclear kinase Dyrk1a in an auto-amplification loop promoting calcineurin/NFAT signalling. Nat Cell Biol *12*, 1220-1227.

Dacheux E, Vincent A, Nazaret N, Combet C, Wierinckx A*, et al.* 2013. BRCA1-dependent translational regulation in breast cancer cells. PLoS One *8*, e67313.

Daniel JM. 2007. Dancing in and out of the nucleus: p120ctn and the transcription factor Kaiso. Biochim Biophys Acta *1773*, 59-68.

Daniel JM, and Reynolds AB. 1999. The catenin p120ctn interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. Mol Cell Biol *19*, 3614-3623.

De La Rosa-Velazquez IA, Rincon-Arano H, Benitez-Bribiesca L, and Recillas-Targa F. 2007. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. Cancer Res *67*, 2577-2585.

De la Torre R, de Sola S, Hernandez G, Farre M*, et al.* 2016. Safety and efficacy of cognitive training plus epigallocatechin-3-gallate in young adults with Down's syndrome TESDAD: a double-blind, randomised, placebo-controlled, phase 2 trial. Lancet Neurol *15*, 801-810.

De la Torre R, De Sola S, Pons M, Duchon A*, et al.* 2014. Epigallocatechin-3-gallate, a DYRK1A inhibitor, rescues cognitive deficits in Down syndrome mouse models and in humans. Mol Nutr Food Res *58*, 278-288.

Defossez PA, Kelly KF, Filion GJ, Perez-Torrado R*, et al.* 2005. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. J Biol Chem *280*, 43017-43023.

del Valle-Perez B, Casagolda D, Lugilde E, Valls G*, et al.* 2016. Wnt controls the transcriptional activity of Kaiso through CK1epsilon-dependent phosphorylation of p120-catenin. J Cell Sci *129*, 873.

Demange L, Abdellah FN, Lozach O, Ferandin Y*, et al.* 2013. Potent inhibitors of CDK5 derived from roscovitine: synthesis, biological evaluation and molecular modelling. Bioorg Med Chem Lett *23*, 125-131.

Deng CX. 2006. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Res *34*, 1416-1426.

Deutsch AJ, Rinner B, Wenzl K, Pichler M*, et al.* 2014. NR4A1-mediated apoptosis suppresses lymphomagenesis and is associated with a favorable cancer-specific survival in patients with aggressive B-cell lymphomas. Blood *123*, 2367-2377.

Dever TE, and Green R. 2012. The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Di Vona C. 2013. Nuclear DYRK1A: new insights into its role within the nucleus. Universitat Pompeu Fabra, Barcelona, Spain.

Di Vona C, Bezdan D, Islam AB, Salichs E*, et al.* 2015. Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. Mol Cell *57*, 506-520.

Dowjat WK, Adayev T, Kuchna I, Nowicki K*, et al.* 2007. Trisomy-driven overexpression of DYRK1A kinase in the brain of subjects with Down syndrome. Neurosci Lett *413*, 77-81.

Duren RP, Boudreaux SP, and Conneely OM. 2016. Genome wide mapping of NR4A binding reveals cooperativity with ETS factors to promote epigenetic activation of distal enhancers in acute myeloid leukemia cells. PLoS One *11*, e0150450.

Durinck S, Moreau Y, Kasprzyk A, Davis S*, et al.* 2005. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics *21*, 3439-3440.

Egan B, Yuan CC, Craske ML, Labhart P*, et al.* 2016. An alternative approach to ChIP-Seq normalization enables detection of genome-wide changes in histone H3 lysine 27 trimethylation upon EZH2 inhibition. PLoS One *11*, e0166438.

Fermi B, Bosio MC, and Dieci G. 2016. Promoter architecture and transcriptional regulation of Abf1-dependent ribosomal protein genes in Saccharomyces cerevisiae. Nucleic Acids Res *44*, 6113-6126.

Ferrari R, Su T, Li B, Bonora G*, et al.* 2012. Reorganization of the host epigenome by a viral oncogene. Genome Res *22*, 1212-1221.

Filion GJ, Zhenilo S, Salozhin S, Yamada D*, et al.* 2006. A family of human zinc finger proteins that bind methylated DNA and repress transcription. Mol Cell Biol *26*, 169-181.

Fotaki V, Dierssen M, Alcantara S, Martinez S*, et al.* 2002. Dyrk1A haploinsufficiency affects viability and causes developmental delay and abnormal brain morphology in mice. Mol Cell Biol *22*, 6636-6647.

Garcia-Cerro S, Martinez P, Vidal V, Corrales A*, et al.* 2014. Overexpression of Dyrk1A is implicated in several cognitive, electrophysiological and neuromorphological alterations found in a mouse model of Down syndrome. PLoS One *9*, e106572.

Garcia-Cerro S, Rueda N, Vidal V, Lantigua S, and Martinez-Cue C. 2017. Normalizing the gene dosage of Dyrk1A in a mouse model of Down syndrome rescues several Alzheimer's disease phenotypes. Neurobiol Dis *106*, 76-88.

Gardini A, Baillat D, Cesaroni M, and Shiekhattar R. 2014. Genome-wide analysis reveals a role for BRCA1 and PALB2 in transcriptional co-activation. EMBO J *33*, 890-905.

Genuario RR, Kelley DE, and Perry RP. 1993. Comparative utilization of transcription factor GABP by the promoters of ribosomal protein genes rpL30 and rpL32. Gene Expr *3*, 279-288.

Gibson DG, Young L, Chuang RY, Venter JC*, et al.* 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.

Giguere V. 2018. Canonical signaling and nuclear activity of mTOR-a teamwork effort to regulate metabolism and cell growth. FEBS J.

Ginzberg MB, Kafri R, and Kirschner M. 2015. Cell biology. On being the right cell size. Science *348*, 1245075.

Glenewinkel F, Cohen MJ, King CR, Kaspar S*, et al.* 2016. The adaptor protein DCAF7 mediates the interaction of the adenovirus E1A oncoprotein with the protein kinases DYRK1A and HIPK2. Sci Rep *6*, 28241.

Gockler N, Jofre G, Papadopoulos C, Soppa U*, et al.* 2009. Harmine specifically inhibits protein kinase DYRK1A and interferes with neurite formation. FEBS J *276*, 6324-6337.

Gorski JJ, Savage KI, Mulligan JM, McDade SS*, et al.* 2011. Profiling of the BRCA1 transcriptome through microarray and ChIP-chip analysis. Nucleic Acids Res *39*, 9536-9548.

Graham FL, and van der Eb AJ. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology *52*, 456-467.

Grant CE, Bailey TL, and Noble WS. 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017-1018.

Grummt I, and Langst G. 2013. Epigenetic control of RNA polymerase I transcription in mammalian cells. Biochim Biophys Acta *1829*, 393-404.

Guedj F, Pereira PL, Najas S, Barallobre MJ*, et al.* 2012. DYRK1A: a master regulatory protein controlling brain growth. Neurobiol Dis *46*, 190-203.

Guedj F, Sebrie C, Rivals I, Ledru A*, et al.* 2009. Green tea polyphenols rescue of brain defects induced by overexpression of DYRK1A. PLoS One *4*, e4606.

Guimera J, Casas C, Estivill X, and Pritchard M. 1999. Human minibrain homologue MNBH/DYRK1: characterization, alternative splicing, differential tissue expression, and overexpression in Down syndrome. Genomics *57*, 407-418.

Guimera J, Casas C, Pucharcos C, Solans A*, et al.* 1996. A human homologue of Drosophila minibrain MNB is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. Hum Mol Genet *5*, 1305-1310.

Guo X, Williams JG, Schug TT, and Li X. 2010. DYRK1A and DYRK3 promote cell survival through phosphorylation and activation of SIRT1. J Biol Chem *285*, 13223-13232.

Gupta V, and Warner JR. 2014. Ribosome-omics of the human ribosome. RNA *20*, 1004-1013.

Gwack Y, Sharma S, Nardone J, Tanasa B*, et al.* 2006. A genome-wide Drosophila RNAi screen identifies DYRK-family kinases as regulators of NFAT. Nature *441*, 646-650.

Hall JM, Lee MK, Newman B, Morrow JE*, et al.* 1990. Linkage of early-onset familial breast cancer to chromosome 17q21. Science *250*, 1684-1689.

Hammerle B, Carnicero A, Elizalde C, Ceron J*, et al.* 2003. Expression patterns and subcellular localization of the Down syndrome candidate protein MNB/DYRK1A suggest a role in late neuronal differentiation. Eur J Neurosci *17*, 2277-2286.

Han J, Miranda-Saavedra D, Luebbering N, Singh A*, et al.* 2012. Deep evolutionary conservation of an intramolecular protein kinase activation mechanism. PLoS One *7*, e29702.

Hanks SK, and Hunter T. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase catalytic domain structure and classification. FASEB J *9*, 576-596.

Hanssen LLP, Kassouf MT, Oudelaar AM, Biggs D*, et al.* 2017. Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. Nat Cell Biol *19*, 952-961.

Harlen KM, and Churchman LS. 2017. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. Nat Rev Mol Cell Biol *18*, 263-273.

Harlen KM, Trotta KL, Smith EE, Mosaheb MM*, et al.* 2016. Comprehensive RNA polymerase II interactomes reveal distinct and varied roles for each phospho-CTD residue. Cell Rep *15*, 2147-2158.

Harte MT, O'Brien GJ, Ryan NM, Gorski JJ*, et al.* 2010. BRD7, a subunit of SWI/SNF complexes, binds directly to BRCA1 and regulates BRCA1-dependent transcription. Cancer Res *70*, 2538-2547.

Hatchi E, Skourti-Stathaki K, Ventz S, Pinello L*, et al.* 2015. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol Cell *57*, 636-647.

Hedrick E, Lee SO, and Safe S. 2017. The nuclear orphan receptor NR4A1 regulates beta1-integrin expression in pancreatic and colon cancer cells and can be targeted by NR4A1 antagonists. Mol Carcinog *56*, 2066-2075.

Hernandez G, Ramirez MJ, Minguillon J, Quiles P*, et al.* 2018. Decapping protein EDC4 regulates DNA repair and phenocopies BRCA1. Nat Commun *9*, 967.

Hille S, Dierck F, Kuhl C, Sosna J*, et al.* 2016. Dyrk1a regulates the cardiomyocyte cell cycle via D-cyclin-dependent Rb/E2f-signalling. Cardiovasc Res *110*, 381-394.

Himpel S, Panzer P, Eirmbter K, Czajkowska H*, et al.* 2001. Identification of the autophosphorylation sites and characterization of their effects in the protein kinase DYRK1A. Biochem J *359*, 497-505.

Himpel S, Tegge W, Frank R, Leder S*, et al.* 2000. Specificity determinants of substrate recognition by the protein kinase DYRK1A. J Biol Chem *275*, 2431-2438.

Hinnebusch AG. 2014. The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem *83*, 779-812.

Hinton CV, Fitzgerald LD, and Thompson ME. 2007. Phosphatidylinositol 3-kinase/Akt signaling enhances nuclear localization and transcriptional activity of BRCA1. Exp Cell Res *313*, 1735-1744.

Hong JY, Park JI, Lee M, Munoz WA*, et al.* 2012. Down's-syndrome-related kinase Dyrk1A modulates the p120-catenin-Kaiso trajectory of the Wnt signaling pathway. J Cell Sci *125*, 561-569.

Hu H, and Li X. 2007. Transcriptional regulation in eukaryotic ribosomal protein genes. Genomics *90*, 421-423.

Hu S, Xie Z, Onishi A, Yu X*, et al.* 2009. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. Cell *139*, 610-622.

Hulsen T, de Vlieg J, and Alkema W. 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics *9*, 488.

Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM*, et al.* 2004. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. Cell *119*, 1041-1054.

Jang SM, Azebi S, Soubigou G, and Muchardt C. 2014. DYRK1A phoshorylates histone H3 to differentially regulate the binding of HP1 isoforms and antagonize HP1-mediated transcriptional repression. EMBO Rep *15*, 686-694.
Jeronimo C, Collin P, and Robert F. 2016. The RNA polymerase II CTD: the increasing complexity of a low-complexity protein domain. J Mol Biol *428*, 2607-2622.
Ji X, Li W, Song J, Wei L, and Liu XS. 2006. CEAS: cis-regulatory element annotation system. Nucleic Acids Res *34*, W551-554.
Johnson N, Cai D, Kennedy RD, Pathania S*, et al.* 2009. Cdk1 participates in BRCA1-dependent S phase checkpoint control in response to DNA damage. Mol Cell *35*, 327-339.
Johnston R, D'Costa Z, Ray S, Gorski J*, et al.* 2016. The identification of a novel role for BRCA1 in regulating RNA polymerase I transcription. Oncotarget *7*, 68097-68110.
Kaczmarski W, Barua M, Mazur-Kolecka B, Frackowiak J*, et al.* 2014. Intracellular distribution of differentially phosphorylated dual-specificity tyrosine phosphorylation-regulated kinase 1A DYRK1A. J Neurosci Res *92*, 162-173.
Kang JE, Choi SA, Park JB, and Chung KC. 2005. Regulation of the proapoptotic activity of huntingtin interacting protein 1 by Dyrk1 and caspase-3 in hippocampal neuroprogenitor cells. J Neurosci Res *81*, 62-72.
Karolchik D, Hinrichs AS, Furey TS, Roskin KM*, et al.* 2004. The UCSC Table Browser data retrieval tool. Nucleic Acids Res *32*, D493-496.
Kehn K, Berro R, Alhaj A, Bottazzi ME*, et al.* 2007. Functional consequences of cyclin D1/BRCA1 interaction in breast cancer cells. Oncogene *26*, 5060-5069.
Kent WJ, Sugnet CW, Furey TS, Roskin KM*, et al.* 2002. The human genome browser at UCSC. Genome Res *12*, 996-1006.
Kentrup H, Becker W, Heukelbach J, Wilmes A*, et al.* 1996. Dyrk, a dual specificity protein kinase with unique structural features whose activity is dependent on tyrosine residues between subdomains VII and VIII. J Biol Chem *271*, 3488-3495.
Khachigian LM. 2018. The Yin and Yang of YY1 in tumor growth and suppression. Int J Cancer. Jan 11. doi: 10.1002/ijc.31255.
Khatter H, Myasnikov AG, Natchiar SK, and Klaholz BP. 2015. Structure of the human 80S ribosome. Nature *520*, 640-645.
Kheradpour P, and Kellis M. 2014. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. Nucleic Acids Res *42*, 2976-2987.
Kida E, Walus M, Jarzabek K, Palminiello S*, et al.* 2011. Form of dual-specificity tyrosine-Y-phosphorylation-regulated kinase 1A nonphosphorylated at tyrosine 145 and 147 is enriched in the nuclei of astroglial cells, adult hippocampal progenitors, and some cholinergic axon terminals. Neuroscience *195*, 112-127.
Kii I, Sumida Y, Goto T, Sonamoto R*, et al.* 2016. Selective inhibition of the kinase DYRK1A by targeting its folding process. Nat Commun *7*, 11391.
Kim D, Won J, Shin DW, Kang J*, et al.* 2004. Regulation of Dyrk1A kinase activity by 14-3-3. Biochem Biophys Res Commun *323*, 499-504.
Kim EJ, Sung JY, Lee HJ, Rhim H*, et al.* 2006a. Dyrk1A phosphorylates alpha-synuclein and enhances intracellular inclusion formation. J Biol Chem *281*, 33250-33257.
Kim H, Lee KS, Kim AK, Choi M*, et al.* 2016a. A chemical with proven clinical safety rescues Down-syndrome-related phenotypes in through DYRK1A inhibition. Dis Model Mech *9*, 839-848.
Kim J, Siverly AN, Chen D, Wang M*, et al.* 2016b. Ablation of miR-10b suppresses oncogene-induced mammary tumorigenesis and metastasis and reactivates tumor-suppressive pathways. Cancer Res *76*, 6424-6435.
Kim JL, Ha GH, Campo L, Denning MF*, et al.* 2017. The role of Rak in the regulation of stability and function of BRCA1. Oncotarget *8*, 86799-86815.
Kim MS, and Hahn JS. 2016. Role of CK2-dependent phosphorylation of Ifh1 and Crf1 in transcriptional regulation of ribosomal protein genes in Saccharomyces cerevisiae. Biochim Biophys Acta *1859*, 1004-1013.
Kim MY, Jeong BC, Lee JH, Kee HJ*, et al.* 2006b. A repressor complex, AP4 transcription factor and geminin, negatively regulates expression of target genes in nonneuronal cells. Proc Natl Acad Sci USA *103*, 13074-13079.
Kim SW, Fang X, Ji H, Paulson AF*, et al.* 2002. Isolation and characterization of XKaiso, a transcriptional repressor that associates with the catenin Xp120ctn in Xenopus laevis. J Biol Chem *277*, 8202-8208.

Kimura R, Kamino K, Yamamoto M, Nuripa A*, et al.* 2007. The DYRK1A gene, encoded in chromosome 21 Down syndrome critical region, bridges between beta-amyloid production and tau phosphorylation in Alzheimer disease. Hum Mol Genet *16*, 15-23.

Kinstrie R, Luebbering N, Miranda-Saavedra D, Sibbet G*, et al.* 2010. Characterization of a domain that transiently converts class 2 DYRKs into intramolecular tyrosine kinases. Sci Signal *3*, ra16.

Kleiman FE, Wu-Baer F, Fonseca D, Kaneko S*, et al.* 2005. BRCA1/BARD1 inhibition of mRNA 3' processing involves targeted degradation of RNA polymerase II. Genes Dev *19*, 1227-1237.

Knight B, Kubik S, Ghosh B, Bruzzone MJ*, et al.* 2014. Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. Genes Dev *28*, 1695-1709.

Kondrashov N, Pusic A, Stumpf CR, Shimizu K*, et al.* 2011. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. Cell *145*, 383-397.

Koromilas AE. 2015. Roles of the translation initiation factor eIF2alpha serine 51 phosphorylation in cancer formation and treatment. Biochim Biophys Acta *1849*, 871-880.

Kress TR, Sabo A, and Amati B. 2015. MYC: connecting selective transcriptional control to global RNA production. Nat Rev Cancer *15*, 593-607.

Kuhn C, Frank D, Will R, Jaschinski C*, et al.* 2009. DYRK1A is a novel negative regulator of cardiomyocyte hypertrophy. J Biol Chem *284*, 17320-17327.

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF*, et al.* 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97.

Kurabayashi N, Hirota T, Sakai M, Sanada K, and Fukada Y. 2010. DYRK1A and glycogen synthase kinase 3beta, a dual-kinase mechanism directing proteasomal degradation of CRY2 for circadian timekeeping. Mol Cell Biol *30*, 1757-1768.

Laguna A, Aranda S, Barallobre MJ, Barhoum R*, et al.* 2008. The protein kinase DYRK1A regulates caspase-9-mediated apoptosis during retina development. Dev Cell *15*, 841-853.

Laguna A, Barallobre MJ, Marchena MA, Mateus C*, et al.* 2013. Triplication of DYRK1A causes retinal structural and functional alterations in Down syndrome. Hum Mol Genet *22*, 2775-2784.

Langmead B, Trapnell C, Pop M, and Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lee HS, Safe S, and Lee SO. 2017. Inactivation of the orphan nuclear receptor NR4A1 contributes to apoptosis induction by fangchinoline in pancreatic cancer cells. Toxicol Appl Pharmacol *332*, 32-39.

Lee SO, Andey T, Jin UH, Kim K*, et al.* 2012. The nuclear receptor TR3 regulates mTORC1 signaling in lung cancer cells expressing wild-type p53. Oncogene *31*, 3265-3276.

Lee Y, Ha J, Kim HJ, Kim YS*, et al.* 2009. Negative feedback Inhibition of NFATc1 by DYRK1A regulates bone homeostasis. J Biol Chem *284*, 33343-33351.

Lempiainen H, and Shore D. 2009. Growth control and ribosome biogenesis. Curr Opin Cell Biol *21*, 855-863.

Lepagnol-Bestel AM, Zvara A, Maussion G, Quignon F*, et al.* 2009. DYRK1A interacts with the REST/NRSF-SWI/SNF chromatin remodelling complex to deregulate gene clusters involved in the neuronal phenotypic traits of Down syndrome. Hum Mol Genet *18*, 1405-1414.

Li D, Jackson RA, Yusoff P, and Guy GR. 2010. Direct association of Sprouty-related protein with an EVH1 domain SPRED 1 or SPRED2 with DYRK1A modifies substrate/kinase interactions. J Biol Chem *285*, 35374-35385.

Li Y, Wang H, Muffat J, Cheng AW*, et al.* 2013. Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. Cell Stem Cell *13*, 446-458.

Litovchick L, Florens LA, Swanson SK, Washburn MP, and DeCaprio JA. 2011. DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. Genes Dev *25*, 801-813.

Liu Q, Liu N, Zang S, Liu H*, et al.* 2014. Tumor suppressor DYRK1A effects on proliferation and chemoresistance of AML cells by downregulating c-Myc. PLoS One *9*, e98853.
Liu Q, Tang Y, Chen L, Liu N*, et al.* 2016. E3 Ligase SCFbetaTrCP-induced DYRK1A protein degradation is essential for cell cycle progression in HEK293 cells. J Biol Chem *291*, 26399-26409.
Liu S, Ginzberg MB, Patel N, Hild M*, et al.* 2018. Size uniformity of animal cells is actively maintained by a p38 MAPK-dependent regulation of G1-length. Elife *7*.
Lochhead PA, Sibbet G, Morrice N, and Cleghon V. 2005. Activation-loop autophosphorylation is mediated by a novel transitional intermediate form of DYRKs. Cell *121*, 925-936.
Loven J, Orlando DA, Sigova AA, Lin CY*, et al.* 2012. Revisiting global gene expression analysis. Cell *151*, 476-482.
Lu M, Zheng L, Han B, Wang L*, et al.* 2011. REST regulates DYRK1A transcription in a negative feedback loop. J Biol Chem *286*, 10755-10763.
Luco SM, Pohl D, Sell E, Wagner JD*, et al.* 2016. Case report of novel DYRK1A mutations in 2 individuals with syndromic intellectual disability and a review of the literature. BMC Med Genet *17*, 15.
Ma X, Zhang K, and Li X. 2009. Evolution of Drosophila ribosomal protein gene core promoters. Gene *432*, 54-59.
Maenz B, Hekerman P, Vela EM, Galceran J, and Becker W. 2008. Characterization of the human DYRK1A promoter and its regulation by the transcription factor E2F1. BMC Mol Biol *9*, 30.
Malinge S, Bliss-Moreau M, Kirsammer G, Diebold L*, et al.* 2012. Increased dosage of the chromosome 21 ortholog Dyrk1a promotes megakaryoblastic leukemia in a murine model of Down syndrome. J Clin Invest *122*, 948-962.
Malygin AA, Parakhnevitch NM, Ivanov AV, Eperon IC, and Karpova GG. 2007. Human ribosomal protein S13 regulates expression of its own gene at the splicing step by a feedback mechanism. Nucleic Acids Res *35*, 6414-6423.
Mao J, Maye P, Kogerman P, Tejedor FJ*, et al.* 2002. Regulation of Gli1 transcriptional activity in the nucleus by Dyrk1. J Biol Chem *277*, 35156-35161.
Marti E, Altafaj X, Dierssen M, de la Luna S*, et al.* 2003. Dyrk1A expression pattern supports specific roles of this kinase in the adult central nervous system. Brain Res *964*, 250-263.
Martin DE, Soulard A, and Hall MN. 2004. TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell *119*, 969-979.
Martinez de Lagran M, Altafaj X, Gallego X, Marti E*, et al.* 2004. Motor phenotypic alterations in TgDyrk1a transgenic mice implicate DYRK1A in Down syndrome motor dysfunction. Neurobiol Dis *15*, 132-142.
Martinez de Lagran M, Benavides-Piccione R, Ballesteros-Yanez I, Calvo M*, et al.* 2012. Dyrk1A influences neuronal morphogenesis through regulation of cytoskeletal dynamics in mammalian cortical neurons. Cereb Cortex *22*, 2867-2877.
Masaki S, Kii I, Sumida Y, Kato-Sumida T*, et al.* 2015. Design and synthesis of a potent inhibitor of class 1 DYRK kinases as a suppressor of adipogenesis. Bioorg Med Chem *23*, 4434-4441.
Maston GA, Evans SK, and Green MR. 2006. Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet *7*, 29-59.
Mazumder B, Sampath P, Seshadri V, Maitra RK*, et al.* 2003. Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. Cell *115*, 187-198.
McElyea SD, Starbuck JM, Tumbleson-Brink DM, Harrington E*, et al.* 2016. Influence of prenatal EGCG treatment and Dyrk1a dosage reduction on craniofacial features associated with Down syndrome. Hum Mol Genet *25*, 4856-4869.
McGary K, and Nudler E. 2013. RNA polymerase and the ribosome: the close relationship. Curr Opin Microbiol *16*, 112-117.
Meine R, Becker W, Falke H, Preu L*, et al.* 2018. Indole-3-carbonitriles as DYRK1A inhibitors by fragment-based drug design. Molecules *23,* 2.
Merkenschlager M, and Nora EP. 2016. CTCF and cohesin in genome folding and transcriptional gene regulation. Annu Rev Genomics Hum Genet *17*, 17-43.
Mills EW, and Green R. 2017. Ribosomopathies: There's strength in numbers. Science *358*.

Moisan A, and Gaudreau L. 2006. The BRCA1 COOH-terminal region acts as an RNA polymerase II carboxyl-terminal domain kinase inhibitor that modulates p21WAF1/CIP1 expression. J Biol Chem *281*, 21119-21130.
Moisan A, Larochelle C, Guillemette B, and Gaudreau L. 2004. BRCA1 can modulate RNA polymerase II carboxy-terminal domain phosphorylation levels. Mol Cell Biol *24*, 6947-6956.
Murakami N, Bolton D, and Hwang YW. 2009. Dyrk1A binds to multiple endocytic proteins required for formation of clathrin-coated vesicles. Biochemistry *48*, 9297-9305.
Myrianthopoulos V, Kritsanida M, Gaboriaud-Kolar N, Magiatis P*, et al.* 2013. Novel inverse binding mode of indirubin derivatives yields improved selectivity for DYRK kinases. ACS Med Chem Lett *4*, 22-26.
Nadeau G, Boufaied N, Moisan A, Lemieux KM*, et al.* 2000. BRCA1 can stimulate gene transcription by a unique mechanism. EMBO Rep *1*, 260-265.
Nagarajan P, Onami TM, Rajagopalan S, Kania S*, et al.* 2009. Role of chromodomain helicase DNA-binding protein 2 in DNA damage response signaling and tumorigenesis. Oncogene *28*, 1053-1062.
Narod SA, and Foulkes WD. 2004. BRCA1 and BRCA2: 1994 and beyond. Nat Rev Cancer *4*, 665-676.
Nepomuceno TC, Fernandes VC, Gomes TT, Carvalho RS*, et al.* 2017. BRCA1 recruitment to damaged DNA sites is dependent on CDK9. Cell Cycle *16*, 665-672.
Nicol JW, Helt GA, Blanchard SG, Jr., Raja A, and Loraine AE. 2009. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics *25*, 2730-2731.
Nojima T, Gomes T, Grosso ARF, Kimura H*, et al.* 2015. Mammalian NET-Seq Reveals Genome-wide nascent transcription coupled to RNA processing. Cell *161*, 526-540.
Nomura M, Gourse R, and Baughman G. 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annu Rev Biochem *53*, 75-117.
O'Leary NA, Wright MW, Brister JR, Ciufo S*, et al.* 2016. Reference sequence RefSeq database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res *44*, D733-745.
Ogawa Y, Nonaka Y, Goto T, Ohnishi E*, et al.* 2010. Development of a novel selective inhibitor of the Down syndrome-related kinase Dyrk1A. Nat Commun *1*, 86.
Orlando DA, Chen MW, Brown VE, Solanki S*, et al.* 2014. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. Cell Rep *9*, 1163-1170.
Ormondroyd E, de la Luna S, and La Thangue NB. 1995. A new member of the DP family, DP-3, with distinct protein products suggests a regulatory role for alternative splicing in the cell cycle transcription factor DRTF1/E2F. Oncogene *11*, 1437-1446.
Ortiz-Abalia J, Sahun I, Altafaj X, Andreu N*, et al.* 2008. Targeting Dyrk1A with AAVshRNA attenuates motor alterations in TgDyrk1A, a mouse model of Down syndrome. Am J Hum Genet *83*, 479-488.
Padovan-Merhar O, Nair GP, Biaesch AG, Mayer A*, et al.* 2015. Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Mol Cell *58*, 339-352.
Park J, Oh Y, Yoo L, Jung MS*, et al.* 2010. Dyrk1A phosphorylates p53 and inhibits proliferation of embryonic neuronal cells. J Biol Chem *285*, 31895-31906.
Park JI, Kim SW, Lyons JP, Ji H*, et al.* 2005. Kaiso/p120-catenin and TCF/beta-catenin complexes coordinately regulate canonical Wnt gene targets. Dev Cell *8*, 843-854.
Pena C, Hurt E, and Panse VG. 2017. Eukaryotic ribosome assembly, transport and quality control. Nat Struct Mol Biol *24*, 689-699.
Perina D, Korolija M, Roller M, Harcet M*, et al.* 2011. Over-represented localized sequence motifs in ribosomal protein gene promoters of basal metazoans. Genomics *98*, 56-63.
Perry RP. 2005. The architecture of mammalian ribosomal protein promoters. BMC Evol Biol *5*, 15.
Pol SU, Polanco JJ, Seidman RA, O'Bara MA*, et al.* 2017. Network-based genomic analysis of human oligodendrocyte progenitor differentiation. Stem Cell Reports *9*, 710-723.
Pozner A, Terooatea TW, and Buck-Koehntop BA. 2016. Cell-specific Kaiso/ZBTB33 regulation of cell cycle through cyclin D1 and cyclin E1. J Biol Chem *291*, 24538-24550.

Pozo N, Zahonero C, Fernandez P, Linares JM*, et al.* 2013. Inhibition of DYRK1A destabilizes EGFR and reduces EGFR-dependent glioblastoma growth. J Clin Invest *123*, 2475-2487.
Proudfoot NJ. 2016. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science *352*, aad9926.
Raaf L, Noll C, Cherifi M, Benazzoug Y*, et al.* 2010. Hyperhomocysteinemia-induced Dyrk1a downregulation results in cardiomyocyte hypertrophy in rats. Int J Cardiol *145*, 306-307.
Rachdi L, Kariyawasam D, Aiello V, Herault Y*, et al.* 2014. Dyrk1A induces pancreatic beta cell mass expansion and improves glucose tolerance. Cell Cycle *13*, 2221-2229.
Radhakrishnan A, Nanjappa V, Raja R, Sathe G*, et al.* 2016. A dual specificity kinase, DYRK1A, as a potential therapeutic target for head and neck squamous cell carcinoma. Sci Rep *6*, 36132.
Raghav SK, Waszak SM, Krier I, Gubelmann C*, et al.* 2012. Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO. Mol Cell *46*, 335-350.
Ramirez F, Ryan DP, Gruning B, Bhardwaj V*, et al.* 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160-165.
Rappsilber J, Mann M, and Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc *2*, 1896-1906.
Raveau M, Shimohata A, Amano K, Miyamoto H, and Yamakawa K. 2018. DYRK1A-haploinsufficiency in mice causes autistic-like features and febrile seizures. Neurobiol Dis *110*, 180-191.
Reid DW, Xu D, Chen P, Yang H, and Sun L. 2017. Integrative analyses of translatome and transcriptome reveal important translational controls in brown and white adipose regulated by microRNAs. Sci Rep *7*, 5681.
Richter JD, and Coller J. 2015. Pausing on polyribosomes: make way for elongation in translational control. Cell *163*, 292-300.
Robinson MD, McCarthy DJ, and Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Robinson MD, and Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol *11*, R25.
Robinson MD, and Smyth GK. 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics *9*, 321-332.
Rodriguez D, Bretones G, Quesada V, Villamor N*, et al.* 2015. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. Blood *126*, 195-202.
Roepcke S, Zhi D, Vingron M, and Arndt PF. 2006. Identification of highly specific localized sequence motifs in human ribosomal protein gene promoters. Gene *365*, 48-56.
Roewenstrunk J. 2016. RNF169 and RNF168 novel substrates of DYRK1A: connecting DYRK1A to DNA-damage repair. University Pompeu Fabra, Barcelona, Spain.
Rogers S, Wells R, and Rechsteiner M. 1986. Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234*, 364-368.
Rosmarin AG, Resendes KK, Yang Z, McMillan JN, and Fleming SL. 2004. GA-binding protein transcription factor: a review of GABP as an integrator of intracellular signaling and protein-protein interactions. Blood Cells Mol Dis *32*, 143-154.
Roy R, Chun J, and Powell SN. 2011. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat Rev Cancer *12*, 68-78.
Ruffner H, Jiang W, Craig AG, Hunter T, and Verma IM. 1999. BRCA1 is phosphorylated at serine 1497 in vivo at a cyclin-dependent kinase 2 phosphorylation site. Mol Cell Biol *19*, 4843-4854.
Ruiz-Velasco M, Kumar M, Lai MC, Bhat P*, et al.* 2017. CTCF-mediated chromatin loops between promoter and gene body regulate alternative splicing across individuals. Cell Syst *5*, 628-637 e626.
Ruvinsky I, Sharon N, Lerer T, Cohen H*, et al.* 2005. Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. Genes Dev *19*, 2199-2211.

Ryoo SR, Cho HJ, Lee HW, Jeong HK*, et al.* 2008. Dual-specificity tyrosineY-phosphorylation regulated kinase 1A-mediated phosphorylation of amyloid precursor protein: evidence for a functional link between Down syndrome and Alzheimer's disease. J Neurochem *104*, 1333-1344.

Ryoo SR, Jeong HK, Radnaabazar C, Yoo JJ*, et al.* 2007. DYRK1A-mediated hyperphosphorylation of Tau. A functional link between Down syndrome and Alzheimer disease. J Biol Chem *282*, 34850-34857.

Saeboe-Larssen S, Lyamouri M, Merriam J, Oksvold MP, and Lambertsson A. 1998. Ribosomal protein insufficiency and the minute syndrome in Drosophila: a dose-response relationship. Genetics *148*, 1215-1224.

Safe S, Imanirad P, Sreevalsan S, Nair V, and Jutooru I. 2014. Transcription factor Sp1, also known as specificity protein 1 as a therapeutic target. Expert Opin Ther Targets *18*, 759-769.

Salichs E. 2008. PolyHistidine repeats and DYRK1A: from the localization to the function. Universitat Pompeu Fabra, Barcelona, Spain.

Salichs E, Ledda A, Mularoni L, Alba MM, and de la Luna S. 2009. Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. PLoS Genet *5*, e1000397.

Sanger F, Nicklen S, and Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA *74*, 5463-5467.

Saunders A, Core LJ, and Lis JT. 2006. Breaking barriers to transcription elongation. Nat Rev Mol Cell Biol *7*, 557-567.

Savage KI, Gorski JJ, Barros EM, Irwin GW*, et al.* 2014. Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. Mol Cell *54*, 445-459.

Saxton RA, and Sabatini DM. 2017. mTOR signaling in growth, metabolism, and disease. Cell *168*, 960-976.

Scales TM, Lin S, Kraus M, Goold RG, and Gordon-Weeks PR. 2009. Nonprimed and DYRK1A-primed GSK3 beta-phosphorylation sites on MAP1B regulate microtubule dynamics in growing axons. J Cell Sci *122*, 2424-2435.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V*, et al.* 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.

Schmoller KM. 2017. The phenomenology of cell size control. Curr Opin Cell Biol *49*, 53-58.

Segales J, Islam AB, Kumar R, Liu QC*, et al.* 2016. Chromatin-wide and transcriptome profiling integration uncovers p38alpha MAPK as a global regulator of skeletal muscle differentiation. Skelet Muscle *6*, 9.

Semba Y, Harada A, Maehara K, Oki S*, et al.* 2017. Chd2 regulates chromatin for proper gene expression toward differentiation in mouse embryonic stem cells. Nucleic Acids Res *45*, 8758-8772.

Shi Z, and Barna M. 2015. Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. Annu Rev Cell Dev Biol *31*, 31-54.

Shi Z, Fujii K, Kovary KM, Genuth NR*, et al.* 2017. Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. Mol Cell *67*, 71-83 e77.

Siggens L, Cordeddu L, Ronnerblad M, Lennartsson A, and Ekwall K. 2015. Transcription-coupled recruitment of human CHD1 and CHD2 influences chromatin accessibility and histone H3 and H3.3 occupancy at active chromatin regions. Epigenetics Chromatin *8*, 4.

Sitz JH, Tigges M, Baumgartel K, Khaspekov LG, and Lutz B. 2004. Dyrk1A potentiates steroid hormone-induced transcription via the chromatin remodeling factor Arip4. Mol Cell Biol *24*, 5821-5834.

Skalska L, Beltran-Nebot M, Ule J, and Jenner RG. 2017. Regulatory feedback from nascent RNA to chromatin and transcription. Nat Rev Mol Cell Biol *18*, 331-337.

Slavov N, Semrau S, Airoldi E, Budnik B, and van Oudenaarden A. 2015. Differential stoichiometry among core ribosomal proteins. Cell Rep *13*, 865-873.

Sonamoto R, Kii I, Koike Y, Sumida Y*, et al.* 2015. Identification of a DYRK1A inhibitor that induces degradation of the target kinase using co-chaperone CDC37 fused with luciferase nanoKAZ. Sci Rep *5*, 12728.

Soppa U, Schumacher J, Florencio Ortiz V, Pasqualon T*, et al.* 2014. The Down syndrome-related protein kinase DYRK1A phosphorylates p27Kip1 and Cyclin D1 and induces cell cycle exit and neuronal differentiation. Cell Cycle *13*, 2084-2100.

Soundararajan M, Roos AK, Savitsky P, Filippakopoulos P*, et al.* 2013. Structures of Down syndrome kinases, DYRKs, reveal mechanisms of kinase activation and substrate recognition. Structure *21*, 986-996.

Soutourina J. 2018. Transcription regulation by the Mediator complex. Nat Rev Mol Cell Biol *19*, 262-274.

Spearman C. 1987. The proof and measurement of association between two things. By C. Spearman, 1904. Am J Psychol *100*, 441-471.

Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H*, et al.* 2003. Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA *9*, 493-501.

Sulima SO, and De Keersmaecker K. 2017. Ribosomal proteins: a novel class of oncogenic drivers. Oncotarget *8*, 89427-89428.

Tahtouh T, Elkins JM, Filippakopoulos P, Soundararajan M*, et al.* 2012. Selectivity, cocrystal structures, and neuroprotective properties of leucettines, a family of protein kinase inhibitors derived from the marine sponge alkaloid leucettamine B. J Med Chem *55*, 9312-9330.

Tee WW, Shen SS, Oksuz O, Narendra V, and Reinberg D. 2014. Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. Cell *156*, 678-690.

Tibbetts RS, Cortez D, Brumbaugh KM, Scully R*, et al.* 2000. Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. Genes Dev *14*, 2989-3002.

Titov DV, Gilman B, He QL, Bhat S*, et al.* 2011. XPB, a subunit of TFIIH, is a target of the natural product triptolide. Nat Chem Biol *7*, 182-188.

Tiwari VK, Stadler MB, Wirbelauer C, Paro R*, et al.* 2011. A chromatin-modifying function of JNK during stem cell differentiation. Nat Genet *44*, 94-100.

Trapnell C, Pachter L, and Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trapnell C, Williams BA, Pertea G, Mortazavi A*, et al.* 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tschop K, Conery AR, Litovchick L, Decaprio JA*, et al.* 2011. A kinase shRNA screen links LATS2 and the pRB tumor suppressor. Genes Dev *25*, 814-830.

Turowski TW, and Tollervey D. 2016. Transcription by RNA polymerase III: insights into mechanism and regulation. Biochem Soc Trans *44*, 1367-1375.

van Riggelen J, Yetil A, and Felsher DW. 2010. MYC as a regulator of ribosome biogenesis and protein synthesis. Nat Rev Cancer *10*, 301-309.

van Roy FM, and McCrea PD. 2005. A role for Kaiso-p120ctn complexes in cancer? Nat Rev Cancer *5*, 956-964.

Vidaki M, Drees F, Saxena T, Lanslots E*, et al.* 2017. A requirement for Mena, an actin regulator, in local mRNA translation in developing neurons. Neuron *95*, 608-622 e605.

Vo Ngoc L, Wang YL, Kassavetis GA, and Kadonaga JT. 2017. The punctilious RNA polymerase II core promoter. Genes Dev *31*, 1289-1301.

von der Heyde S, Fromm-Dornieden C, Salinas-Riester G, Beissbarth T, and Baumgartner BG. 2014. Dynamics of mRNA and polysomal abundance in early 3T3-L1 adipogenesis. BMC Genomics *15*, 381.

Walte A, Ruben K, Birner-Gruenberger R, Preisinger C*, et al.* 2013. Mechanism of dual specificity kinase activity of DYRK1A. FEBS J *280*, 4495-4511.

Wang J, Zhuang J, Iyer S, Lin X*, et al.* 2012. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. Genome Res *22*, 1798-1812.

Wang P, Wang L, Chen L, and Sun X. 2017. Dual-specificity tyrosine-phosphorylation regulated kinase 1A Gene Transcription is regulated by Myocyte Enhancer Factor 2D. Sci Rep *7*, 7240.

Wang RH, He JP, Su ML, Luo J*, et al.* 2013. The orphan receptor TR3 participates in angiotensin II-induced cardiac hypertrophy by controlling mTOR signalling. EMBO Mol Med *5*, 137-148.

Wang YL, Duttke SH, Chen K, Johnston J*, et al.* 2014. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. Genes Dev *28*, 1550-1555.
Wegiel J, Kuchna I, Nowicki K, Frackowiak J*, et al.* 2004. Cell type- and brain structure-specific patterns of distribution of minibrain kinase in human brain. Brain Res *1010*, 69-80.
Wickham H 2009. ggplot2: Elegant graphics for data analysis. J Stat Software *35*, Book review 1.
Wong QW, Li J, Ng SR, Lim SG*, et al.* 2014. RPL39L is an example of a recently evolved ribosomal protein paralog that shows highly specific tissue expression patterns and is upregulated in ESCs and HCC tumors. RNA Biol *11*, 33-41.
Woods YL, Cohen P, Becker W, Jakes R*, et al.* 2001. The kinase DYRK phosphorylates protein-synthesis initiation factor eIF2Bepsilon at Ser539 and the microtubule-associated protein tau at Thr212: potential role for DYRK as a glycogen synthase kinase 3-priming kinase. Biochem J *355*, 609-615.
Wu H, Bi J, Peng Y, Huo L*, et al.* 2017. Nuclear receptor NR4A1 is a tumor suppressor down-regulated in triple-negative breast cancer. Oncotarget *8*, 54364-54377.
Wu J, Lu LY, and Yu X. 2010. The role of BRCA1 in DNA damage response. Protein Cell *1*, 117-123.
Wyrwicz LS, Gaj P, Hoffmann M, Rychlewski L, and Ostrowski J. 2007. A common cis-element in promoters of protein synthesis and cell cycle genes. Acta Biochim Pol *54*, 89-98.
Xue S, and Barna M. 2012. Specialized ribosomes: a new frontier in gene regulation and organismal biology. Nat Rev Mol Cell Biol *13*, 355-369.
Xue S, Tian S, Fujii K, Kladwang W*, et al.* 2015. RNA regulons in Hox 5' UTRs confer ribosome specificity to gene regulation. Nature *517*, 33-38.
Yabut O, Domogauer J, and D'Arcangelo G. 2010. Dyrk1A overexpression inhibits proliferation and induces premature neuronal differentiation of neural progenitor cells. J Neurosci *30*, 4004-4014.
Yamashita D, Sano Y, Adachi Y, Okamoto Y*, et al.* 2007. hDREF regulates cell proliferation and expression of ribosomal protein genes. Mol Cell Biol *27*, 2003-2013.
Yang CS, Wang X, Lu G, and Picinich SC. 2009. Cancer prevention by tea: animal studies, molecular mechanisms and human relevance. Nat Rev Cancer *9*, 429-439.
Yarden RI, and Brody LC. 1999. BRCA1 interacts with components of the histone deacetylase complex. Proc Natl Acad Sci USA *96*, 4983-4988.
Yoganathan T, Bhat NK, and Sells BH. 1992. A positive regulator of the ribosomal protein gene, beta factor, belongs to the ETS oncoprotein family. Biochem J *287 Pt 2*, 349-353.
Yoon HG, Chan DW, Reynolds AB, Qin J, and Wong J. 2003. N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. Mol Cell *12*, 723-734.
Yoshihama M, Uechi T, Asakawa S, Kawasaki K*, et al.* 2002. The human ribosomal protein genes: sequencing and comparative analysis of 73 genes. Genome Res *12*, 379-390.
Zawel L, and Reinberg D. 1995. Common themes in assembly and function of eukaryotic transcription complexes. Annu Rev Biochem *64*, 533-561.
Zehavi Y, Kedmi A, Ideses D, and Juven-Gershon T. 2015. TRF2: TRansForming the view of general transcription factors. Transcription *6*, 1-6.
Zhang J, Willers H, Feng Z, Ghosh JC*, et al.* 2004. Chk2 phosphorylation of BRCA1 regulates DNA double-strand break repair. Mol Cell Biol *24*, 708-718.
Zhang X, Chiang HC, Wang Y, Zhang C*, et al.* 2017. Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. Nat Commun *8*, 15908.
Zhang Y, Duc AC, Rao S, Sun XL*, et al.* 2013. Control of hematopoietic stem cell emergence by antagonistic functions of ribosomal protein paralogs. Dev Cell *24*, 411-425.
Zhang Y, Liao JM, Zeng SX, and Lu H. 2011. p53 downregulates Down syndrome-associated DYRK1A through miR-1246. EMBO Rep *12*, 811-817.
Zhao C, Deng Y, Liu L, Yu K*, et al.* 2016. Dual regulatory switch through interactions of Tcf7l2/Tcf4 with stage-specific partners propels oligodendroglial maturation. Nat Commun *7*, 10883.

Zufferey R, Nagy D, Mandel RJ, Naldini L, and Trono D. 1997. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol *15*, 871-875.

Annexes

1. ANNEX I

Nomenclature for proteins from the large ribosomal subunit

New name[#]	Bacteria name	Yeast name	Human name
uL1	L1	L1	L10A
uL2	L2	L2	L8
uL3	L3	L3	L3
uL4	L4	L4	L4
uL5	L5	L11	L11
uL6	L6	L9	L9
eL6	–	L6	L6
eL8	–	L8	L7A
bL9	L9	–	–
uL10	L10	P0	P0
uL11	L11	L12	L12
bL12	L7/L12	–	–
uL13	L13	L16	L13A
eL13	–	L13	L13
uL14	L14	L23	L23
eL14	–	L14	L14
uL15	L15	L28	L27A
eL15	–	L15	L15
uL16	L16	L10	L10
bL17	L17	–	–
uL18	L18	L5	L5
eL18	–	L18	L18
bL19	L19	–	–
eL19	–	L19	L19
bL20	L20	–	–
eL20	–	L20	L18A
bL21	L21	–	–
eL21	–	L21	L21
uL22	L22	L17	L17
eL22	–	L22	L22

uL23	L23	L25	L23A
uL24	L24	L26	L26
eL24	–	L24	L24
bL25	L25	–	–
bL27	L27	–	–
eL27	–	L27	L27
bL28	L28	–	–
eL28	–	–	L28
uL29	L29	L35	L35
eL29	–	L29	L29
uL30	L30	L7	L7
eL30	–	L30	L30
bL31	L31	–	–
eL31	–	L31	L31
bL32	L32	–	–
eL32	–	L32	L32
bL33	L33	–	–
eL33	–	L33	L35A
bL34	L34	–	–
eL34	–	L34	L34
bL35	L35	–	–
bL36	L36	–	–
eL36	–	L36	L36
eL37	–	L37	L37
eL38	–	L38	L38
eL39	–	L39	L39
eL40	–	L40	L40
eL41	–	L41	L41
eL42	–	L42	L36A
eL43	–	L43	L37A
P1/P2	–	P1/P2 (AB)	P1/P2 (αβ)

New nomenclature for proteins from the small ribosomal subunit

New name[#]	Bacteria name	Yeast name	Human name
bS1	S1	–	–
eS1	–	S1	S3A
uS2	S2	S0	SA
uS3	S3	S3	S3
uS4	S4	S9	S9
eS4	–	S4	S4
uS5	S5	S2	S2
bS6	S6	–	–
eS6	–	S6	S6
uS7	S7	S5	S5
eS7	–	S7	S7
uS8	S8	S22	S15A
eS8	–	S8	S8
uS9	S9	S16	S16
uS10	S10	S20	S20
eS10	–	S10	S10
uS11	S11	S14	S14
uS12	S12	S23	S23
eS12	–	S12	S12
uS13	S13	S18	S18
uS14	S14	S29	S29
uS15	S15	S13	S13
bS16	S16	–	–
uS17	S17	S11	S11
eS17	–	S17	S17
bS18	S18	–	–
uS19	S19	S15	S15
eS19	–	S19	S19
bS20	S20	–	–
bS21	S21	–	–
bTHX	THX	–	–
eS21	–	S21	S21

eS24	–	S24	S24
eS25	–	S25	S25
eS26	–	S26	S26
eS27	–	S27	S27
eS28	–	S28	S28
eS30	–	S30	S30
eS31	–	S31	S27A
RACK1	–	Asc1	RACK1

Laura Barba was supported by a FPU predoctoral fellowship of the Spanish Ministery of Education, Culture and Sport-MECD (FPU13/02400). This work was supported by the Spanish Ministry of Economy and Competitiveness-MINECO (BFU2013-44513-P and BFU2016-76141-P) and the Secretariat of Universities and Research-Government of Catalonia (2009SGR1464). The group belongs to the Rare Diseases Networking Biomedical Center (CIBERER). The CRG is a 'Centro de Excelencia Severo Ochoa 2013-2017'

Agradecimientos

Me gustaría que estas líneas sirvieran para expresar mi más profundo y sincero agradecimiento a todas aquellas personas que con su ayuda han colaborado en la realización del presente trabajo, ya que sin ellas no habría sido posible. En primer lugar, a Susana, gracias por darme la oportunidad de llevar a cabo mi tesis doctoral en tu laboratorio, ha supuesto todo un reto y un aprendizaje para mí, tanto a nivel profesional como personal. Me gustaría agradecerte tu constante confianza en mi trabajo, incluso en aquellos momentos en los que ningún experimento funcionaba. También por todas aquellas largas reuniones en base a las cuales considero que he ido adquiriendo un buen criterio científico. Muchas gracias. En segundo lugar, a Chiara, mi ChIP-partner, has sido una persona esencial en este proyecto. No hace falta que te diga que sin ti esta tesis no estaría escrita hoy. Me has ayudado desde el primer día hasta el último. No sólo me has abierto al mundo de la cromatina, sino que también has contribuido a mi desarrollo como científica. Además, muchos días he podido sobrellevar las frustraciones que conlleva este trabajo gracias a ti. En especial, mil gracias por tu esfuerzo y paciencia en estos últimos meses de escritura haciendo millones de análisis y aguantando mis agobios. Spero che tu sia fiera di me! Grazie mille. Gracias de corazón.

Mi más sincera gratitud al resto de compañeros del 671 por ser tan estupendos. Sin vuestros ánimos y ese buen ambiente que creáis esta tesis habría sido mucho más difícil de sobrellevar: a mis queridas mamis, Alicia y Kriszti, muchas gracias por haberme ayudado siempre que os lo he pedido. Ali, porque hemos podido compartir expresiones del castellano profundo que tenías olvidadas entre tanto “guiri”. Kriszti, por ser mi chica mutagénesis: köszönöm!; a Julia, porque, aunque te hemos desquiciado con nuestro desorden, has sabido tener paciencia y como buena enfermera me has cuidado siempre que lo he necesitado; a Jacopo, por tener siempre una sonrisa en la cara y por no volverte loco entre tantas chicas. Te mando mis ánimos de cara a la escritura de la tesis y estoy 100% segura de que todo va a ir genial. ¡Vas a triunfar con tu superclon!; to Rianne, thanks for helping me whenever I needed a favor. I encourage you to fight in your last year of the PhD. You can do it!; a Francesco Sottile, porque, aunque tu paso por el 671 ha sido breve, ya en ese poco tiempo me has transmitido tus ánimos e ideas y me has ayudado con algunos experimentos; a Karen, nuestra chilena adoptiva, también gracias por tu apoyo y buena onda; y a Borja, aún no hemos tenido tiempo de conocernos, pero me da la sensación de que vas a encajar muy bien en el lab, pese a que seas demasiado ordenado.

Me gustaría dar las gracias a Roberto Ferrari, no solo por tu paciencia analizando y reanalizando tantas ChIPs, sino también por responder a todas mis preguntas y por todos tus consejos; a Enrique Blanco, por el análisis del RNA-Seq y también por ejercer de profesor conmigo y llenarme de optimismo siempre que nos hemos cruzado por el pasillo; a Marc Talló y al resto de compañeros del 368, por ayudarme siempre que lo he necesitado con la máquina medieval de los polisomas; a Gwendal DuJardin y a Marta Inglés, por vuestra infinita paciencia con los EMSAs; y a Lara de Llobet, por tu consideración, apoyo y ánimos desde el primer día que nos conocimos. En general, gracias a todos nuestros vecinos del lab de Fátima Gebauer, por estar siempre dispuestos a ayudarme siempre que lo he necesitado. Grazie mille! Moltes gracièss! Merci beaucoup!

También me gustaría expresar mi agradecimiento a los servicios de Genómica, Bioinformática y Proteómica del PRBB, así como a los servicios de mantenimiento y soporte informático porque sin su predisposición el proyecto habría sido mucho más difícil de llevar a cabo. In addition, I would like to thank all the people from the GRSC program that have helped me in any occasion along the PhD as well as my Thesis Committee members, Luciano di Croce, Bill Keyes and Pura Muñoz, for your advice and assistance.

Por otro lado, no podía dejar de dar las gracias a aquellas personas fuera del CRG que, aún sin entender ni una palabra relacionada con este proyecto, han estado apoyándome en todo momento: a mis grupos de yoga y sevillanas, porque el ambiente que hemos creado juntos durante las clases me ha servido como vía de escape muchas veces. A mis amigos de Vilanova, por haberme acogido en este maravilloso pueblo. Por supuesto, gracias a todos mis amigos de Madrid, por no haber perdido el contacto conmigo pese a la distancia: a Fermín, por armarse de paciencia y atreverse con conceptos de biología para el diseño de esta portada; a mis queridos Clonos, a Marco y a Marina, porque hemos compartido penas y alegrías durante la carrera y el máster y las hemos seguido compartiendo durante el doctorado; a Isa, gracias por esas conversaciones por Skype pese a que cada vez te encuentres más lejos; a mis agapornis de colores, por encontrar siempre un huequito para ver al pequeño kiwi; a Vero y a Lore, por bajar de la sierra madrileña tan solo para poder verme un rato; y a Montse, no solo porque tienes el récord de visitas sino por estar ahí siempre que lo he necesitado.

Por último, mi más sincero agradecimiento a mi familia, especialmente a mi madre y hermano, por aguantarme en mis malos momentos y mandarme la fuerza necesaria para ir superando todos los obstáculos y

poder llegar hasta aquí, sin vosotros no lo habría conseguido; y a Fer, porque te embarcaste conmigo en esta aventura sin pensártelo dos veces y por hacerme sonreir todos los días. Os quiero. Gracias de corazón.