Chromosoma
DOI 10.1007/s00412-014-0503-8

REVIEW

The Bucentaur (BCNT) protein family: a long-neglected class of essential proteins required for chromatin/chromosome organization and function

Giovanni Messina · Emanuele Celauro · Maria Teresa Atterrato · Ennio Giordano · Shintaro Iwashita · Patrizio Dimitri

Received: 27 October 2014 / Revised: 5 December 2014 / Accepted: 5 December 2014

Abstract The evolutionarily conserved Bucentaur (BCNT) protein superfamily was identified about two decades ago in bovines, but its biological role has long remained largely unknown. Sparse studies in the literature suggest that BCNT proteins perform important functions during development. Only recently, a functional analysis of the *Drosophila* BCNT ortholog, called YETI, has provided evidence that it is essential for proper fly development and plays roles in chromatin organization. Here, we introduce the BCNT proteins and comprehensively review data that contribute to clarify their function and mechanistic clues on how they may control development in multicellular organisms.

Keywords Bucentaur (BCNT) · YETI · ATP-dependent chromatin remodeling complexes · H2A variant · Epigenetics

G. Messina · E. Celauro · M. T. Atterrato · P. Dimitri (✉)
Dipartimento di Biologia e Biotecnologie "Charles Darwin", Sapienza Università di Roma, Rome, Italy
e-mail: patrizio.dimitri@gmail.com

P. Dimitri
e-mail: patrizio.dimitri@uniroma1.it

G. Messina · E. Celauro · M. T. Atterrato · P. Dimitri
Istituto Pasteur Fondazione Cenci-Bolognetti, Sapienza Università di Roma, Rome, Italy

E. Giordano
Dipartimento di Biologia, Università Federico II, Naples, Italy

S. Iwashita
Faculty of Pharmacy, Iwaki Meisei University, Iwaki, Japan

Published online: 31 December 2014

Introduction

Our primary obligation is not following, as animals do, the crowd of those who precede us
Lucio Anneo Seneca

The man who follows the crowd will usually get no farther than the crowd
Albert Einstein

Conformism is frequent in Science, and, as a consequence, potentially important facts and results can be ignored or secreted for long time. This appears to be the case for the evolutionarily conserved BCNT protein superfamily which, in spite of its widespread distribution across both animal and plant kingdoms, has long been neglected and its function(s) remained largely unexplored.

The BCNT proteins in ruminants: the origins

The first member of the BCNT protein superfamily, p97BCNT/CFDP2, was discovered in bovine brain during screening for hybridoma producing monoclonal antibodies against a Ras GTPase (Ras GAP) activating protein (Nobukuni et al. 1997; Iwashita and Osada 2011). A purified glutathione-S-transferase (GST) fusion protein of rat Rasa2 (GAP1m) was used to immunize BALB/c mice and the generated hybridomas were screened by Western blotting using bovine brain extract. A single band with a molecular mass of 97 kDa was found in five independent clones, exactly the expected size of rat Rasa2. However, the 97-kDa protein purified from bovine brain extracts by affinity chromatography turned out to be different from Rasa2, as shown by amino acid sequence analysis after protease-digestion. The identified 97-kDa protein not only shared traces of amino acid sequence of Ras GAP

but also contained the amino acid sequence corresponding to the boundary of GST-tag and Ras GAP fusion, constructed to prepare the immune antigen. It consisted of an acidic N-terminal region, a retrotransposon-derived 325-amino acid region (the RTE domain) and of two 40-amino acid intrarepeat (IR) units. The RTE domain was 72 % identical to that of a ruminant specific retrotransposon, called Bov-B LINE. The corresponding clone carrying an open reading frame encoding a polypeptide of 592 amino acids was named Bcnt, after Bucentaur, a mythological creature that is half man and half ox. The 97-kDa protein was named p97Bcnt.

The *Bcnt* genes

Sequence analysis of the bovine genome detected three tandemly aligned *Bcnt* genes: *Bcnt/Cfdp1* and *p97Bcnt/Cfdp2* and *p97Bcnt2/Bcnt2* (Fig. 1). The *p97Bcnt/Cfdp2* and *p97Bcnt2/Bcnt2* paralogs resulted from two different tandem duplication events (Iwashita et al. 2009; Iwashita and Osada 2011). In the first event, the ancestral *Bcnt* gene underwent partial duplication followed by RTE-1 targeting; subsequently, the duplicated copy, *p97Bcnt/Cfdp2*, underwent further duplication giving rise to *p97Bcnt2/BCNT2*. As the result, the two *Bcnt* paralogs recruited the AP-endonuclease-like domain of RTE-1 in the middle of molecule.

The three *Bcnt* genes are located in the bovine chromosome 18, spanning about 180 kb, a region with synteny to human chromosome 16q23 and mouse chromosome 8 (Fig. 1). The cluster is proximally flanked by the breast cancer antiestrogen resistance 1 (*Bcar1*) gene and distally by the transmembrane protein 170A (*Tmem170A*) gene, an organization which is conserved not only in mammals but also in frog and zebrafish. Noteworthy, at least in human and in mouse, the distal region corresponds to a complex locus named *Tmem170* and *Cfdp1* (Thierry-Mieg and Thierry-Mieg 2006, www.ncbi.nlm.nih.gov/IEB/Research/Acembly/). In ruminants, both ancestral *Bcnt/Cfdp1* and its paralogous genes are present, while other animals have only *Bcnt/Cfdp1*.

The BCNT proteins

The BCNT protein superfamily shows a widespread distribution in animals and plants and is classified into Pfam 07572 in the Pfam database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=pfam07572).

The BCNT/CFDP1 bovine protein shares homologous N-terminal acidic regions and a single IR unit of 40 amino acids, but lacks the RTE domain (Fig. 1; Iwashita et al. 2003, 2009). BCNT/CFDP1 contains a highly conserved 82-amino acid region at the C-terminus, termed BCNT-C or BCNT domain, which is absent in p97BCNT/CFDP2 and in p97BCNT2 paralogs. A specific characteristic of the three BCNT-related proteins in bovine is given by their different numbers of IR (intramolecular repeat) units: BCNT/CFDP1, p97BCNT/CFDP2, and p97BCNT2 carry one, two, or three IR units, respectively. The p97BCNT2/BCNT2 protein is highly homologous to p97BCNT/CFDP2 and comprises an acidic N-terminal region, a 324-amino acid RTE domain, and three IR units instead of the two found in p97BCNT/CFDP2 at the C-terminal region. Whereas the BCNT-C portion is highly conserved among almost all eukaryotes, the N-terminal regions are less conserved (see examples in Fig. 2). For example, the amino acids of the *Drosophila* BCNT (YETI) are ~50 % identical to those of bovine BCNT/CFDP1 in the C-terminal region, while the N-terminal region shows only ~22 % identity.

Notably, both BCNT/CFDP1 and p97BCNT/CFDP2 are phosphoproteins potentially phosphorylated in vitro on serine residues by the casein kinase II (Iwashita et al. 1999). Recently, the two phosphorylated serine residues in human BCNT, 116S in the N-terminal region and 250S in the C-terminal region, were identified by mass spectrometric analysis (Dephoure et al. 2008). The ~175-amino acid N-terminal portion of the three BCNT-

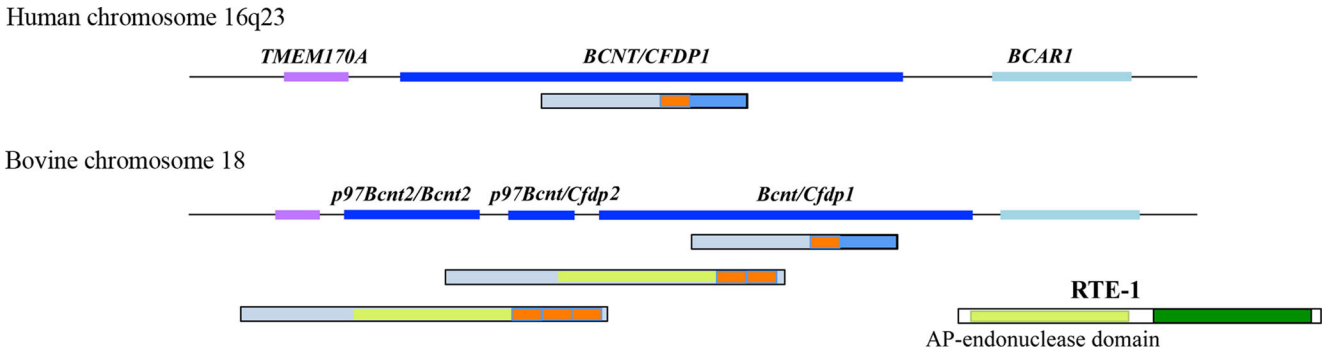


Fig. 1 Organization of the *Bcnt* locus and architecture of BCNT proteins in humans and bovines. The scheme shows a comparison between the human *Cfdp1* region in chromosome 16q23 and the corresponding region of bovine chromosome 18 with the three tandemly aligned *Bcnt* genes (*Bcnt/Cfdp1*, *p97Bcnt/Cfdp2*, and *p97Bcnt2/Bcnt2*). In both human and bovine genomes, *Bcnt* genes (*blue*) are proximally flanked by the breast cancer antiestrogen resistance 1 (*BCAR1*) gene (*pale blue*) and distally by the transmembrane protein 170A (*Tmem170A*) gene (*violet*). The architecture of the corresponding BCNT proteins is also shown in different colors: BCNT domain (*blue*), IR repeats (*orange*), AP-endonuclease domain (*yellow*)

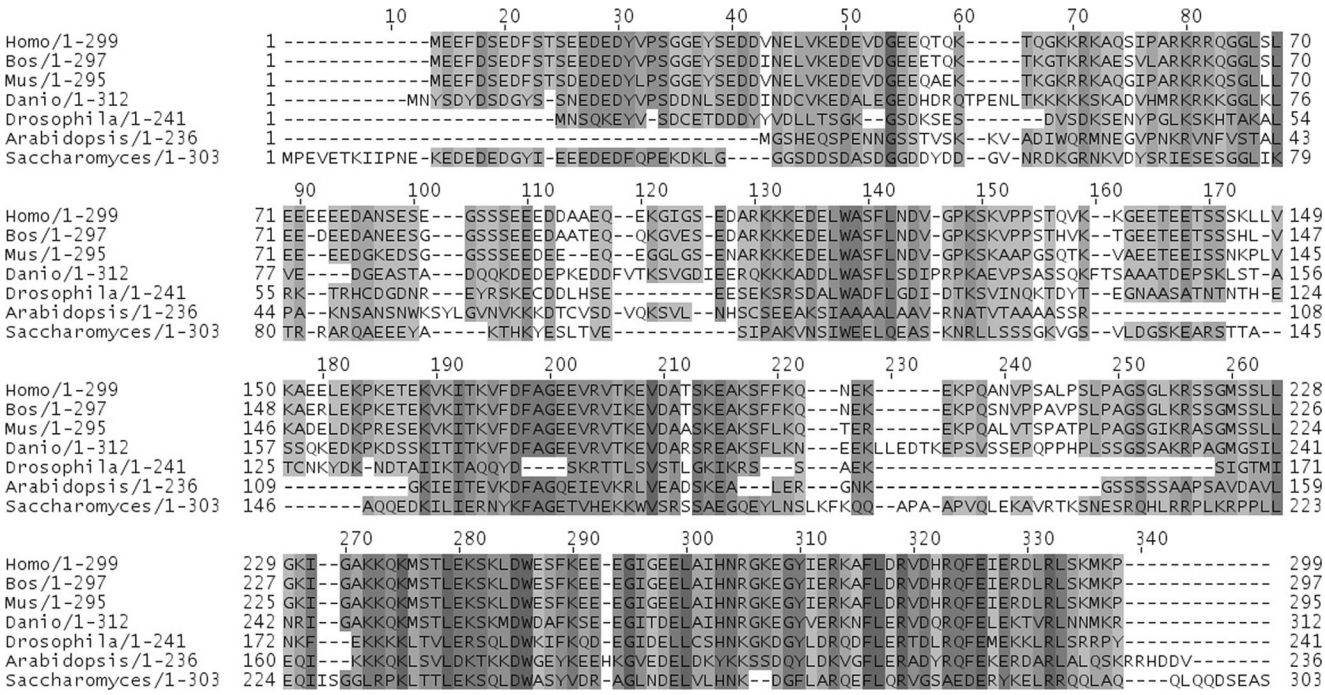


Fig. 2 The BCNT proteins. Blast alignments of BCNT protein sequences from *Homo sapiens*, *Bos taurus*, *Mus musculus*, *Danio rerio*, *Drosophila melanogaster*, *Arabidopsis thaliana*, and *Saccharomyces cerevisiae*. Note the high conservation of protein sequence (about 45 % similarity) in the last 80 residues of the C-terminal end. In *Drosophila melanogaster*, a recessive lethal allele of the *Bcnt gene Yeti*, called EMS31, carries a single base pair substitution (from A to T) at position 316, where codon AAA for K is changed to a TAA stop codon (Messina et al. 2014)

related proteins is acidic as a whole but contains several arginine/lysine-rich elements, including a putative nuclear targeting motif of Arg-Lys-Arg-Lys (~61–64th).

The three BCNT-related proteins in bovine were found in both the cytosolic and nuclear fractions of MDBK cells; in addition, both p97BCNT/CFDP2 and p97BCNT2 were detected in the chromatin fractions (Iwashita et al. 1999, 2009; Iwashita and Osada 2011). Moreover, a GFP-rat BCNT fusion protein expressed in primary rat astrocytes exclusively localized in nuclei (Nakashima and Song, unpublished data). These results suggested that BCNT family members may potentially function as shuttle molecules between the cytosol and nuclei.

The biological role(s) of BCNT protein superfamily

Despite its widespread distribution in animals and plants, the function(s) of BCNT superfamily is largely unknown (Iwashita et al. 2003). Studies conducted in different organisms suggested that BCNT proteins perform essential functions during development. In particular, sparse evidence indicate that BCNT proteins in vertebrates may be implicated in craniofacial development and osteogenesis. Here, we report a summary of what is known in the literature about BCNT orthologs in *Saccharomyces cerevisiae* (SWC5), *Drosophila melanogaster* (YETI), chicken (CENP-29), zebrafish (RLTPR), *Mus musculus* (CP27), and *Homo sapiens* (CFDP1). Sequence alignments of these BCNT members are shown in Fig. 2.

SWC5, the *yeast* BCNT ortholog

The SWC5 protein (also called YBR231C/AOR1) is a component of the SWR1 ATP-dependent chromatin remodeling complex (Wu et al. 2009). The purified SWR1 complex specifically catalyzes displacement of histone H2A from chromatin and its replacement with the H2A variant, Htz1 (Kobor et al. 2004; Mizuguchi et al. 2004). The SWR1 complex contains up to 14 distinct subunits, some of which (SWR1, SWC2 SWC3, SWC5, SWV6, SWC7, ARP6, YAF9, and BDF1) are encoded by genes that are nonessential for cell viability. Other subunits of the complex (ACT1, ARP4, SWC4, RVB1, and RVB2) are essential. The SWR1 ATPase subunit provides a platform for assembly of the entire SWR1 complex.

The SWR1-mediated Htz1 exchange is a multistep reaction consisting of the assembly of the SWR1 complex, its binding to Htz1 substrate, followed by H2A displacement and loading of Htz1 into the vacant site (Wu et al. 2005); it is likely that the units of histone exchange are H2A-H2B and Htz1-H2B dimers rather than Htz1 and H2A monomers. In vitro studies reported a lack of histone exchange in Swc5 deleted mutants, indicating that Swc5 is required for the replacement reaction (Wu et al. 2005). Since the binding of SWR1 complex to Htz1

is not dependent on Swc5, the latter may be required in separate step(s) of the histone–exchange reaction after Htz1 binding (Wu et al. 2005). In vivo results support this conclusion, showing that the absence of Swc5 nearly eliminated the amount of Htz1 bound to chromatin and also suggesting that Swc5 is required for the binding and destabilization of the H2A/H2B dimer (Morillo-Huesca et al. 2010).

The SWC5 protein, however, is not essential, but *Swc5* null mutants are defective in chromosome/plasmid maintenance and show a decrease in resistance to hydroxyurea (*Saccharomyces* Genome Database) and increased heat sensitivity. Notably, changes in the subcellular localization of SWC5 and other SWI/SNF proteins, were observed in that it shuttles between nuclei and cytoplasm in response to oxygen concentration (Dastidar et al. 2012). Thus, the composition of a chromatin remodeling complex can undergo dramatical changes depending upon environmental conditions. Furthermore, the molecular and proteomic analyses of genetically and phenotypically diverse strains with distinct stress sensitivities, indicates that Swc5 may play a role in response to oxidative stress (Codlin et al. 2013).

YETI, the essential BCNT ortholog of *Drosophila*

The *Drosophila melanogaster* BCNT member, YETI, is a 241-aa-long protein encoded by *Yeti*, previously called CG40218, a 967-bp-long gene located in the heterochromatin of chromosome 2 (Corradini et al. 2003; Dimitri et al. 2003; Wisniewski et al. 2003; Messina et al. 2014). The functions of the YETI protein have been investigated in the last decade. YETI was originally defined as a kinesin-binding protein able to bind both subunits of the microtubule-based motor kinesin-I (Wisniewski et al. 2003). Immunolocalization experiments showed both a nuclear and cytoplasmic localization of epitope-tagged YETI in *Drosophila* S2 cultured cells (Wisniewski et al. 2003). In those experiments, however, it remained unclear whether YETI is located within the nuclei or associated with the nuclear membrane.

Recent findings shed light on the functions of YETI and provided the first functional evidence showing that a BCNT protein is essential for proper development and individual viability (Messina et al. 2014). YETI is found in cell nuclei and is present at multiple sites on polytene chromosomes during *Drosophila* development (Fig. 3a, b); its depletion causes lethality before pupation and defects in higher-order chromatin organization (Fig. 3c, d). These defects are accompanied by severe impairment in the association of histone H2A.V, nucleosomal histones, and epigenetic marks on polytene chromosomes. We have also evidence indicating that YETI binds to chromatin through its conserved BCNT domain and interacts with the histone variant H2A.V, Domino-A (DOM-A), and HP1a (Messina et al. 2014). DOM-A ATPase is the ortholog of the yeast Swr1 protein and belongs to the *Drosophila* DOM/dTip60 ATP-dependent chromatin-remodeling complex (Kusch et al. 2004). The DOM/Tip60 complex includes dozen subunits with homology to those of SWR1 and NuA4 complexes in yeast (reviewed by March-Díaz and Reyes 2009). Other stable subunits of the complex are BAP55, GAS41, PONTIN, REPTIN, NIPPED-A, E(Pc), YL1,DMA, ACT87B, MRG15, MRGB, TRA1, ING3, and EAF6(Reviewed by March-Díaz and Reyes 2009). In vitro, the DOM/Tip60 complex catalyzes the exchange of phosphorylated H2A.V with H2A during the extinction of DNA damage (Kusch et al. 2004) and is also proposed to exchange back H2A with H2A.V (Baldi and Becker 2013; Messina et al. 2014).

Together, these results strongly suggest that YETI contributes to the exchange of the variant H2A.V onto nucleosomes. YETI may conceivably perform its function as a subunit of the DOM/Tip60 complex (Fig. 4), similarly to the yeast homologue SWC5. The dramatic defect in H2A deposition also found in *Yeti* mutants suggests that H2A is removed by the DOM/Tip60 complex, but not replaced with H2A.V in YETI-defective cells, resulting in depletion of both histones and thus disrupting chromatin integrity and function (Messina et al. 2014). A similar scenario has been proposed in yeast (Morillo-Huesca et al. 2010). Alternatively, YETI may be a H2A.V chaperone. A number of H2A-H2B chaperones are known that participate in control of transcription initiation and chromatin modifications (Avvakumov et al. 2011). The delivery of H2A.Z in yeast is also assisted by Chz1, a specific chaperone that is displaced upon H2A.Z-H2B binding (Luk et al. 2007). Nap1, in addition to binding canonical H2A-H2B histones, can also interact with H2AZ-H2B variant dimers (Mizuguchi et al. 2004; Kobor et al. 2004).

Independent experimental evidence show that YETI also interacts with HP1a in *Drosophila melanogaster* (Messina et al. 2014; Ryu et al. 2014) thus YETI may be a multifaceted chromatin protein, also mediating the targeting of HP1a to chromatin remodeling regions. The loss of YETI-HP1a interaction may contribute to the global chromatin unbalance found in YETI-depleted cells (Messina et al. 2014).

CP27, the BCNT mouse ortholog

The *cp27* gene in mouse was isolated and cloned from a mouse E11 lgt11 library using an antibody against Cem1, a tooth cementum-related protein which is a mark of a distinct expression pattern of mouse craniofacial development (Diekwisch et al. 1999). The cp27 gene contains an open reading frame of 888 nucleotides. The putative amino acid sequence contains 295 amino acids, and its calculated molecular mass is 33 kDa. On Western Blots, however, a polyclonal antibody against CP27 protein detected a single epitope at 27 kDa, for this reason and for its expression pattern, the gene was named cp27 (craniofacial protein 27). Sequence analysis

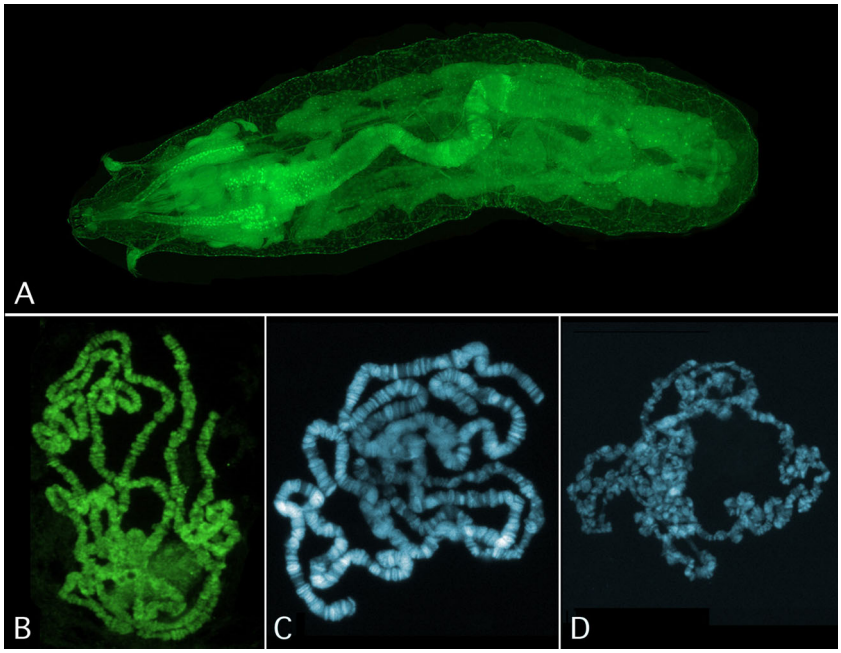


Fig. 3 YETI is a nuclear protein, and its loss causes chromatin defects. **a** Third instar larva of *Drosophila melanogaster* expressing UAS-*Yeti-GFP* transgene under control of *Actin-GAL* 4. YETI-GFP fusion protein is highly abundant in the nuclei. **b** On salivary gland polytene chromosomes, YETI-GFP maps to multiple sites on euchromatic arms and to the chromocenter. **c** DAPI-stained polytene chromosomes from wild-type Oregon-R strain. **d** DAPI-stained polytene chromosomes of *Yeti* mutant larva. Chromosome organization in *Yeti* mutant genotypes is clearly perturbed

of the hypothetical CP27 protein showed that it is an acid protein, and structure predictions suggested that CP27 contains a helix–loop–helix configuration including a total of ten α-helices with alternating β-turns.

The CP27 protein showed a localization at the epithelial-mesenchymal boundary of the developing organs, including the teeth, salivary glands, heart, lung, and liver (Diekwisch et al. 1999). Based on these results, the authors proposed that CP27 may be involved in organogenesis and craniofacial development, but no direct experimental evidence supporting these conclusions were provided. Later on, the CP27 protein was suggested to be an extracellular matrix protein implicated in tooth development (Diekwisch et al. 2002; Diekwisch and Luan 2002).

More recently, experimental evidence suggested that changes in *cp27* gene expression can be functionally relevant during differentiation occurring in mouse development. First, the *cp27* gene was found to be upregulated during experimentally induced osteodifferentiation and is therefore proposed to be required for proper bone development (Bustos-Valenzuela et al. 2011). Second, cleft palate induction by persistent expression of PAX 3 gene in neural crest was found to be correlated with *cp27* downregulation (Wu et al. 2008). In addition, unpublished results indicate that CP27 is a subunit of the mouse SRCAP chromatin remodeling complexes (T. Diekwisch, personal communication), in accord with the roles of SWC5 and YETI proteins in chromatin remodeling.

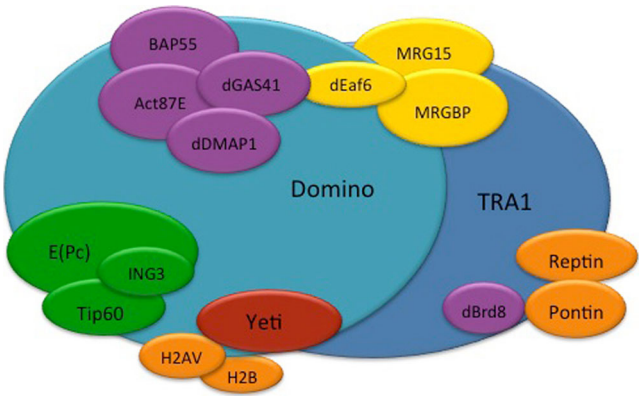


Fig. 4 The *Drosophila melanogaster* DOM/Tip60 chromatin remodeling complex. Cartoon showing the chromatin remodeling complex with its known subunits (modified from van Attikum and Gasser 2005). The YETI protein is shown in *red*

CFDP1, the human BCNT ortholog

The CFDP1 protein is encoded by the *Cfdp1* (craniofacial development protein 1) gene. *Cfdp1* is 140-kb long with seven exons and six introns and maps to 16 in 16q22.2-q22.3 and is flanked proximally by *BCAR1* and distally by the *TMEM170A* (Fig. 1). By microarrays, the *Cfdp1* expression is detectable in a wide range of human tissues, including cancer tissues.

Cfdp1 is predicted to undergo alternative splicing, giving rise to two diverse mRNAs which differ for the presence of an exon encoding the BCNT domain. Using a monoclonal antibody against the CFDP1 protein, two different bands of about 50 and 37 kDa were indeed found in HeLa cells (Fig. 5a; G. Messina, MT. Atterrato, L. Piacentini, A. Losada, and P. Dimitri, unpublished). By IF, we also found that CFDP1 is enriched into the nuclei of HeLa cells (Fig. 5b, c; G. Messina, M.T. Atterrato, L. Piacentini, A. Losada, and P. Dimitri, unpublished).

The *Cfdp1* gene maps in proximity to several loci associated with inherited craniofacial disease genes (Diekwisch

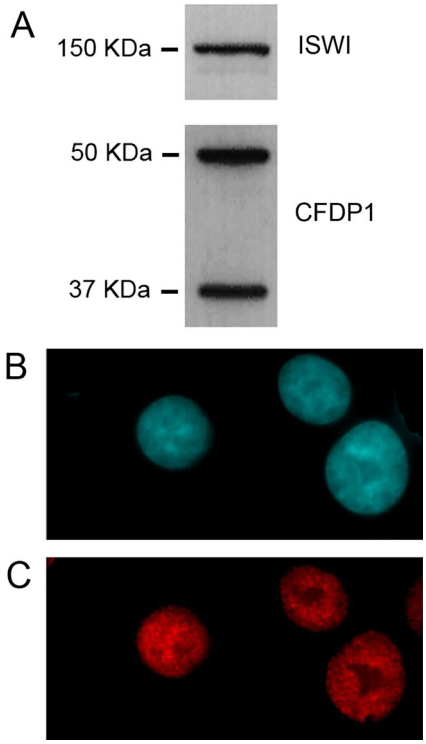


Fig. 5 CFDP1 in HeLa cells. **a** By Western Blotting, a monoclonal antibody against CFDP1 detected two sharp bands of about 50 and 37 kDa. It is possible that these bands correspond to two isoforms of the protein produced by alternative splicing. ISWI, the highly conserved member of the SWI2/SNF2 family of ATPases, was used as loading control. **b** DAPI staining of HeLa cells is shown in *blue*, **c** anti-CFDP1 immunostaining shown in *red*

et al. 1999) and such association have been considered suggestive for an involvement of this gene in human craniofacial development. However, thus far, there are no experimental data proving such an involvement, but only indirect suggestions. In particular, *Cfdp1* was indeed shown to be a target of TFII-I transcription factors, which are in turn encoded by genes suggested to be primary candidate genes responsible for craniofacial abnormalities and other defects associated with Williams-Beuren disease (Makeyev and Bayarsaihan 2011). Notably, the *BCAR1-CFDP1-TMEM170A* locus has been also suggested to be a determinant of carotid intima–media thickness and coronary artery disease risk (Gertow et al. 2012).

Given that *yeast* SWC5 and *Drosophila* YETI belong to the Swir1 and Tip60 chromatin remodeling complexes, respectively, it is possible that CFDP1 is a subunit of the human p400/Tip60 and/or SRCAP chromatin remodeling complexes, which share a dozen of subunits with the corresponding yeast and *Drosophila* complexes (reviewed by Clapier and Cairns 2009). Although the function of CFDP1 in human cells is still unclear, its involvement in chromatin organization is suggested by its nuclear localization (Fig. 5b, c). In addition, by integrative global proteomic profiling approach, that coupled chromatographic separation of human cell extracts and quantitative tandem mass spectrometry, Havugimana et al. (2012) found that CFDP1 interacts with members of the SRCAP chromatin remodeling complex. By Co-IP assays, we found that CFDP1 interacts with Arp6, a know subunit of both p400/Tip60 and SRCAP complexes (G. Messina, unpublished). Thus, together, the experimental evidence indicate that CFDP1 plays a role in chromatin remodeling.

In a systematic search for protein–protein interactions in human cells, Rual et al. (2005) found evidence of interaction of CFDP1 with SMAD3 and Ewing's sarcoma (EWS) proteins, two factors which are both involved in transcriptional regulation. SMAD3 is a member of the SMAD protein family, which are intracellular signaling components of the TGF-β superfamily. SMAD3 forms a heterodymeric complex with SMAD4 which allows its translocation into the nucleus where it acts as a modulator of transcriptional processes (Inman and Hill 2002). EWS is a member of the TET (TLS/EWS/TAF15) family of RNA- and DNA-binding proteins whose expression is altered in cancer (Adam et al. 2013). A frequent chromosomal translocation called t(11;22) (q24;q12) fuses the EWS gene on chromosome 22 to the FLI1 gene on chromosome 11 and causes the Ewing sarcoma and other tumors of the Ewing family (Riggi and Stamenkovic 2007). Mutations of EWSR1 gene have also been associated to amyotrophic lateral sclerosis (Couthouis et al. 2012).

What about BCNT proteins in other vertebrates?

Among other vertebrates, data on BCNT orthologs are only available in birds and fishes. In a recent study that integrated quantitative proteomics with bioinformatic analysis, the BCNT member of chicken, called CENP-29, has been reported to be associated with kinetocores and therefore was suggested to be involved in chromosome segregation (Ohta et al. 2010).

In *Danio rerio*, the BCNT gene called *rltpr*, was found to be developmentally expressed in the eye, CNS and in the branchial arches which lead to the craniofacial structures (Thisse and Thisse 2004). Our unpublished data show that at 72 h postfertilization, the *rltpr* transcript is abundant in the cephalic region of the embryo, mainly in the forebrain area, midbrain up to the midbrain/hindbrain boundary, optic tectum, and presumptive branchial arches (Fig. 6; E. Celauro, A. Rodriguez, A. Raya Chamorro, and P. Dimitri).

Intrinsically disordered proteins, BCNT, and chromatin

Intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) represent structurally ill-defined proteins and protein segments. IDPs and proteins carrying IDRs

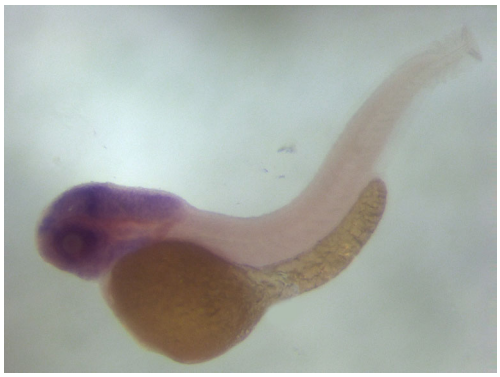


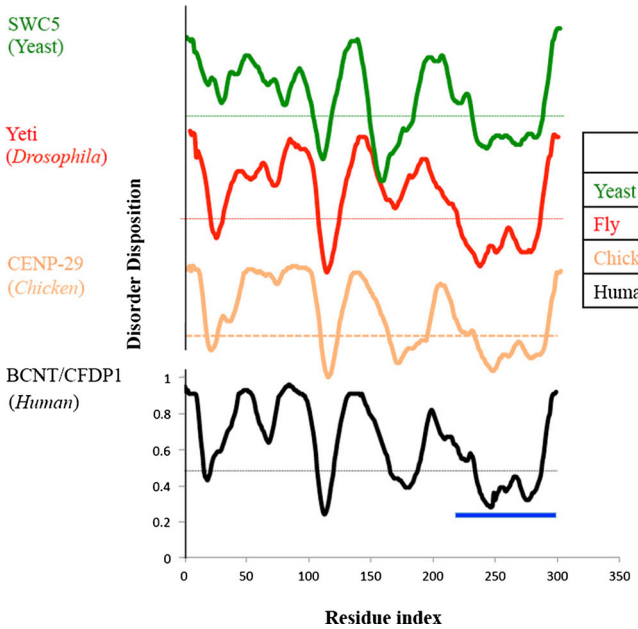
Fig. 6 The *rltpr* gene expression in *Danio rerio* embryo. Whole mount in situ hybridization of *rltpr* mRNA riboprobe on wild-type embryos. At larval protruding-mouth stage (72 h postfertilization), the *rltpr* transcript is restricted to the anterior part of the forming body

lack a fixed or ordered three-dimensional structure or are partially structured with multiple domains connected by flexible linkers. They are abundant in the eukaryotic proteomes and have numerous biological activities (reviewed by Fuxreiter et al. 2008). The discovery of IDPs has challenged the traditional concept of the Anfinsen dogma (Anfinsen 1973), where protein function depends on a fixed three-dimensional structure (Oldfield and Dunker 2014).

Proteins involved in transcriptional control and chromatin regulation have been found to be intrinsically disordered or to carry IDRs critical for chromatin remodeling processes (reviewed by Fuxreiter et al. 2008). IDPs and IDRs are present in N-terminal histone tails and are involved in the interactions with non-histone proteins and multiprotein complexes (Hansen et al. 1998). For example, the N-terminal domain of the core histone H4 functions as an assembly platform for different chromatin remodeling complexes (Clapier et al. 2001; Xiao et al. 2001). Moreover, IDPs also seem to have important roles in the function of ATP-dependent chromatin remodeling complexes, such as ISWI, CHRAC, NURF, RSC, and SWI/SNF, which regulate accessibility of the genomic DNA. It has been hypothesized that IDPs allow malleability in the structure of chromatin remodeling and modifying complexes, which in turn facilitates their interaction with the equally structurally malleable chromatin fiber (Fuxreiter et al. 2008).

Among ID chromatin proteins, the leukemia fusion target AF9, an intrinsically disordered transcriptional regulator, is one of the fruitful examples (Leach et al. 2013). It has been clearly shown that a 79-aa region at the C-terminal AF9, normally a random coil, forms a mixed α-helix and β-structure upon incubation with 14-aa peptide of AF4, one of its known binding partners. The data suggests that whereas electrostatic interaction between AF9 and AF4 determines the binding specificity, extensive hydrophobic interaction can affect the folding .

The BCNT/CFDP1 protein in ruminants is entirely disordered, and its two paralogs p97BCNT/CFDP2 and p97BCNT2, which lack the BCNT-C domain, also contain largely disordered regions, except for a highly ordered AP-endonuclease-like domain derived from the RTE-1 element



	Length (aa)	BCNT-C domain
Yeast	303	214-286
Fly	241	159-239
Chicken	290	212-288
Human	299	220-297

Fig. 7 Prediction of BCNT proteins disordered regions. Amino acid sequences of *Saccharomyces cerevisiae* SWR1-complex protein 5 (Swc 5) (accession # P38326), *Drosophila* Yeti (# EED86089), chicken CENP-29 (#Q75QI0), and human BCNT/CFDP1 (#Q9UEE9) analyzed by the PrDOS software (a protein disorder prediction system; Ishida and Kinoshita 2007). The probability values were converted to obtain a specific disordered disposition profile for each protein. The horizontal axis (residue index) was normalized to the length of 303 a.a. to compare the relative patterns. A disorder disposition higher than 0.5 in more than 50 % of the protein indicates a disorder propensity

(Fig. 1; Iwashita and Osada 2011). A remarkable feature that distinguishes the three ruminant BCNT members is represented by a different number of IR units: BCNT/CFDP1 has a single IR, while p97BCNT/CFDP2 and p97BCNT2 carry two and three IR units, respectively (Fig. 1). Although the aa sequences of the N-terminal region and of the IR units are quite variable among different BCNT proteins (Iwashita et al. 2003; Messina et al. 2014), their predicted plots of disorder profiles are quite similar (Fig. 7).

A critical issue is how ID chromatin remodelers can acquire chromatin binding activity (Tayal *et al.* 2014). It has been well recognized that post-translational modifications (PTMs), such as phosphorylation and acetylation (Oldfield and Dunker 2014) may play on-and-off switches leading to dimerization of IDPs or IDRs and to complex formation. For example, interaction between AF9 and AF4 (see before) is also regulated by phosphorylation of T766 of AF4, which caused the ~30-fold reduction in their affinity (Leach et al. 2000).

Mammalian BCNT has been supposed to be regulated by PTMs. Indeed, human and bovine BCNTs were found to be substrates for in vitro acetylation by the co-transcriptional activator CBP/p300, resulting in the acetylation of at least 15 lysine residues (Iwashita et al. 2013). Three of them were found in a recombinant human BCNT expressed in HEK 293 cells, only when pre-cultured with deacetylase inhibitors. In addition, seven phosphorylated sites were found in the exogenous human His-tag BCNT-C, including the DWESF (Iwashita et al. 2014). Phosphorylation of these sites may also trigger the mammalian BCNT chromatin binding activity. Phosphorylation of SMAD family members is indeed known to induce shuttling activity of these proteins between the cytoplasmic and nuclear compartments via dimerization or cellular complex formation (Moustakas et al. 2011). Intriguingly, a SMAD protein is supposed to interact with human CFDP1 (see previous paragraphs).

Conclusions

In this review, we have tried to illustrate an updated picture of the experimental data available thus far on the BCNT protein family. The emerging scenario indicates that the function of BCNT proteins is conserved throughout evolution of eukaryotic cells, in that they play important roles in chromatin organization and function. In particular, yeast, *Drosophila*, and human BCNT proteins belong to evolutionary conserved chromatin remodeling complexes that exhibit multifaceted biological roles. For example, the Tip60 complex is involved in a wide variety of cellular functions, including transcriptional regulation, DNA repair, cell cycle progression, chromosome stability, stem cell maintenance, and differentiation (Sapountzi et al. 2006; Tea and Luo 2011). It is also worth noting that mutations in genes encoding proteins directly or potentially implicated in chromatin dynamics are crucial players in cancer and in human developmental disorders (Bickmore and van der Maarel 2003; Fog et al. 2012; Harmacek et al. 2014). Thus, we believe that future studies aimed at further exploring the mechanisms by which BCNT protein family governs chromatin state and controls development in model organisms will expand our understanding of mechanisms underpinning cellular differentiation, human pathologies, and developmental diseases.

Acknowledgments The research of PD laboratory was supported by grants from Istituto Pasteur-Fondazione Cenci-Bolognetti and Fondazione Roma - Terzo Settore. The authors thank Dr. Si-Young Song, Master Kentaro Nakashima, and Dr. Yoshiko Ohno-Iwashita for the helpful discussion, and for Professor Tomoyasu Inoue for advice in preparing Fig. 7.

References

Adam J, Yang M, Soga T, Pollard PJ (2013) Rare insights into cancer biology. Oncogene 33:2547–2556

Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181:223–230

Avvakumov N, Nourani A, Cote J (2011) Histone chaperones: modulators of chromatin marks. Mol Cell 41:502–514

Baldi S, Becker PB (2013) The variant histone H2A.V of Drosophila—three roles, two guises. Chromosoma 122:245–258

Bickmore WA, van der Maarel SM (2003) Perturbations of chromatin structure in human genetic disease: recent advances. Hum Mol Genet 2:207–213

Bustos-Valenzuela JC, Fujita A, Halcsik E, Granjeiro JM and Sogayar MC (2011) Unveiling novel genes upregulated by both rhBMP2 and rhBMP7 during early osteoblastic transdifferentiation of C2C12 cells. BMC Res Notes 4:370

Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. Annu Rev Biochem 78:273–304

Clapier CR, Langst G, Corona DF, Becker PB, Nightingale KP (2001) Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. Mol Cell Biol 21:875–883

Codlin S, Marsellach-Castellv FX, Clement-Zizab M, Papadakisc M, Schmidtd A et al (2013) A systems approach linking genotype and environment to phenotype: oxidative stress response mechanisms in fission yeast EMBO Conference on Fission Yeast: pombe 7th International Fission Yeast Meeting

Corradini N, Rossi F, Verni F, Dimitri P (2003) FISH analysis of Drosophila melanogaster heterochromatin using BACs and P elements. Chromosoma 112:26–37

Couthouis J, Hart MP, Erion R, King OD, Diaz Z et al (2012) Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. Hum Mol Genet 21:2899–28911

Dastidar RG, Hooda J, Shah A, Cao TM, Henke RM, Zhang L (2012) The nuclear localization of SWI/SNF proteins is subjected to oxygen regulation. Cell Biosci 2:30. doi:10.1186/2045-3701-2-30.

Dephoure N, Zhou C, Villén J, Beausoleil SA, Bakalarski CE, Elledge SJ, Gygi SP (2008) A quantitative atlas of mitotic phosphorylation. Proc Natl Acad Sci U S A 105:10762–10767. doi:10.1073/pnas.0805139105

Diekwisch TG, Marches F, Williams A, Luan X (1999) Cloning, gene expression, and characterization of CP27, a novel gene in mouse embryogenesis. Gene 235:19–30

Diekwisch TG, Luan X (2002) CP27 function is necessary for cell survival and differentiation during tooth morphogenesis in organ culture. Gene 287:141–147.

Diekwisch TG, Luan X, McIntosh JE (2002) CP27 localization in the dental lamina basement membrane and in the stellate reticulum of developing teeth. J Histochem Cytochem 50:583–586

Dimitri P, Corradini N, Rossi F, Vernì F, Cenci G et al (2003) Vital genes in the heterochromatin of chromosomes 2 and 3 of Drosophila melanogaster. Genetica 117:209–215

Fog CK, Galli GG, Lund AH (2012) PRDM proteins: important players in differentiation and disease. Bioessays 34:50–60

Fuxreiter M, Tompa P, Simon I, Uversky VN, Hansen JC et al (2008) Malleable machines take shape in eukaryotic transcriptional regulation. Nat Chem Biol 4:728–737

Gertow K, Sennblad B, Strawbridge RJ, Ohrvik J, Zabaneh D et al (2012) Identification of the BCAR1-CFDP1-TMEM170A locus as a determinant of carotid intima-media thickness and coronary artery disease risk. Circ Cardiovasc Genet 5:656–665

Hansen JC, Tse C, Wolffe AP (1998) Structure and function of the core histone N-termini: more than meets the eye. Biochemistry 37: 17637–17641

Harmacek L, Watkins-Chow DE, Chen J, Jones KL, Pavan WJ et al (2014) A unique missense allele of BAF155, a core BAF chromatin remodeling complex protein, causes neural tube closure defects in mice. Dev Neurobiol 74:483–497

Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL et al (2012) A census of human soluble protein complexes. Cell 150:1068–1081

Inman GJ, Hill CS (2002) Stoichiometry of active smad-transcription factor complexes on DNA. J Biol Chem 277:51008–51016

Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence, Nucleic Acids Res, 35, Web Server issue

Iwashita S, Osada N (2011) Bucentaur (Bcnt) gene family: gene duplication and tetrotransposon insertion, chapter 21 of Gene duplication, pp 383–400, Felix Friedberg (Ed.) InTech (Rijeka, Croatia) ISBN: 978-953-307-387-3, DOI:10.5772/22351. http://www.intechopen.com/books/gene-duplication/bucentaur-bcnt-gene-family-gene-duplication-and-retrotransposon-insertion

Iwashita S, Nobukuni T, Tanaka S, Kobayashi M, Iwanaga T et al (1999) Partial nuclear localization of a bovine phosphoprotein, BCNT, that includes a region derived from a LINE repetitive sequence in Ruminantia. Biochim Biophys Acta 1427:408–416

Iwashita S, Osada N, Itoh T, Sezaki M, Oshima K et al (2003) A transposable element-mediated gene divergence that directly produces a novel type bovine Bcnt protein including the endonuclease domain of RTE-1. Mol Biol Evol 20:1556–1563

Iwashita S, Nakashima K, Sasaki M, Osada N, Song S-Y (2009) Multiple duplication of the bucentaur gene family, which recruits the APE-like domain of retrotransposon: identification of a novel homolog and distinct cellular expression. Gene 435:88–95

Iwashita S, Yasuda T, Suzuki T, Saotome A, Nakashima K et al (2013) Molecular heterogeneity of Bucentaur (Bcnt)/Cfdp1, a potential component of chromatin remodeling complex—isoforms and post-translational acetylation. 86th Annual meeting of Japanese Biochemical Society (Yokohama)

Iwashita S, Suzuki T, Nakashima K, Kobayashi T, Yasuda T et al (2014) Molecular basis for heterogeneity of phosphorylated Bcnt/Cfdp1, an evolutionarily conserved chromatin organization factor 87th Annual meeting of Japanese Biochemical Society (Kyoto)

Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin J, Jennings JL et al (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol 2:E131

Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL et al (2004) Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. Science 306:2084–2087

Leach BI, Kuntimaddi A, Schmidt CR, Cierpicki T, Johnson SA et al (2013) Leukemia fusion target AF9 is an intrinsically disordered transcriptional regulator that recruits multiple partners via coupled folding and binding. Structure 21:176–183

Leach TJ, Mazzeo M, Chotkowski HL, Madigan JP, Wotring MG, Glaser RL (2000) Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. J Biol Chem 275: 23267–23272

Luk E, Vu ND, Patteson K, Mizuguchi G, Wu WH, Ranjan A, Backus J, Sen S, Lewis M, Bai Y, Wu C (2007) Mol Cell 25:357–368.

Makeyev AV, Bayarsaihan D (2011) Molecular basis of Williams-Beuren syndrome: TFII-I regulated targets involved in craniofacial development. Cleft Palate Craniofac J 48:109–116

March-Díaz R, Reyes JC (2009) The beauty of being a variant: H2A.Z and the SWR1 complex in plants. Mol Plant 2:565–577

Messina G, Damia E, Fanti L, Atterrato MT, Celauro E et al (2014) *Yeti*, an essential *Drosophila melanogaster* gene, encodes a protein required for chromatin organization. J Cell Sci 127:2577–2588

Mizuguchi G, Shen X, Landry J, Wu WH, Sen S et al (2004) ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science 303:343–348

Morillo-Huesca M, Clemente-Ruiz M, Andújar E, Prado F (2010) The SWR1 histone replacement complex causes genetic instability and genome-wide transcription misregulation in the absence of H2A.Z. PLoS ONE 5:e12143

Moustakas A, Serhiy Souchelnytskyi S, Heldin C-H (2011) Smad regulation in TGF-β signal transduction. J Cell Sci 114:4359–4369

Nobukuni T, Kobayashi M, Omori A, Ichinose S, Iwanaga T et al (1997) An Alu-linked repetitive sequence corresponding to 280 amino acids is expressed in a novel bovine protein, but not in its human homologue. J Biol Chem 272:2801–2807

Ohta S, Bukowski-Wills JC, Sanchez-Pulido L, Alves FL, Wood L et al (2010) The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. Cell 142: 810–821

Oldfield CJ, Dunker AK (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. Annu Rev Biochem 83: 553–584

Riggi N, Stamenkovic I (2007) The biology of Ewing sarcoma. Cancer Lett 254:1–10

Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A et al (2005) Towards a proteome-scale map of the human protein-protein interaction network. Nature 437:1173–1178

Ryu H-W, Lee DH, Florens L, Swanson SK, Washburn MP et al (2014) Analysis of the heterochromatin protein 1 (HP1) interactome in Drosophila. J Proteome 102:137–147

Sapountzi V, Logan IR, Robson CR (2006) Cellular functions of TIP60. Int J Biochem Cell Biol 38:1496–1509

Tayal N, Choudhary P, Pandit SB, Sandhu KS (2014) Evolutionarily conserved and conformationally constrained short peptides might serve as DNA recognition elements in intrinsically disordered regions. Mol BioSyst 10:1469–1480

Tea JS, Luo L (2011) The chromatin remodeling factor Bap55 functions through the TIP60 complex to regulate olfactory projection neuron dendrite targeting. Neural Dev 6:5. doi:10.1186/1749-8104-6-5

Thierry-Mieg D, Thierry-Mieg J (2006) AceView: a comprehensive cDNA-supported gene and transcripts annotation Genome Biol (Suppl 1):S12

Thisse B, Thisse C (2004) Fast released clones: high-throughput expression analysis. ZFIN

van Attikum H, Gasser SM (2005) The histone code at DNA breaks: a guide to repair? Nat Rev Mol Cell Biol 6:757–765

Wisniewski TP, Tanzi CL, Gindhart JG (2003) The Drosophila kinesin-I associated protein YETI binds both kinesin subunits. Biol Cell 95: 595–602

Wu W-H, Alami S, Luk E, Wu C-H, Sen S et al (2005) Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. Nat Struct Mol Biol 12:1064–1071

Wu M, Li J, Engleka KA, Zhou B, Lu MM et al (2008) Persistent expression of Pax3 in the neural crest causes cleft palate and defective osteogenesis in mice. J Clin Invest 118:2076–2087

Wu W-H, Wu C-H, Ladurner A, Mizuguchi G, Wei D et al (2009) N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complexes. J Biol Chem 284:6200–6207

Xiao H, Sandaltzopoulos R, Wang HM, Hamiche A, Ranallo R et al (2001) Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. Mol Cell 8:531–543