perspectives

The rapid development of "omics" technologies (genomics, transcriptomics, proteomics, metabolomics, and others) has allowed for a more detailed understanding of complex biological systems. However, analytical approaches for meaningful interpretation of multilayer omics datasets are lagging behind the technological advancements. The sparse, amorphous, distributed, and poorly reproducible state of omics datasets and the lack of standard for generating such data in many disciplines further complicate the analysis of these datasets (Ioannidis, 2005). These bottlenecks pose the importance of developing stringent computational strategies and a validation regime that can distinguish between true signals and noise (Ioannidis and Khoury, 2011). Interdisciplinary approaches are required to bridge the growing gap between technological development, biomedical research, and computational biology. Thus, converging loops between theory and experiment should help to understand the dynamics of biological networks and processes. Integrative analyses of different and multifaceted biological datasets should facilitate the study of human genetic disorders.

In this Outlook, I firstly discuss the scientific rationale for carrying out multi-disciplinary research on a rare human genetic disorder and further outline how that can benefit society. I then describe and provide an overview on some of the key mechanistic insights unravelled by my colleagues and myself through the course of our studies. Finally, some directions which could enhance the evolution of the field of systems biology are discussed.

Remarkable insights from a rare event

DNA sequence polymorphisms contribute to individual differences in disease susceptibility. As genetic information can be passed onto mRNA and proteins that perform cellular functions, genetic studies have often focused on the one-to-one relationship between phenotype and genomic variation as the basis for knowledge discovery. There are, however, common patterns that underlie the diversity, complexity, development and progression of genetic diseases. Hence, looking for shared molecular activities, across organisms or processes with similar characteristics (such as ageing, late-onset neuromuscular and protein aggregation disorders), can lead to uncovering new insights into mechanisms that are important for diverse biological conditions. In particular, such combined strategies would facilitate the study of rare diseases. There are an estimated 8,000 rare disorders, many of which are known to be of genetic origin (Stolk et al., 2006; Schieppati et al., 2008). Given the low prevalence of rare diseases, it is particularly difficult to employ traditional approaches and, therefore, they require special integrative efforts to improve discovery of underlying mechanisms. Notably, in spite of the low prevalence of each rare disease, about 30

million people are estimated to be affected by a rare disease in Europe (Kaplan and Laing, 2004). In the work presented in this thesis, my colleagues and I have mainly focused on developing computational approaches and conducting an interdisciplinary study to unravel novel associations based on shared functional features. Here, I discuss an outlook on how such strategies can lead to significant findings that benefit society. I do this based on the result of our investigations on Oculopharyngeal muscular dystrophy (OPMD).

PABPN1, the protein mutated in OPMD, regulates poly(A) tail length and RNA stability (Lemay et al., 2010; Kuhn et al., 2009). As such, PABPN1 plays an important role in a variety of cellular processes (Kuhn et al., 2009; Wahle, 1991; Wahle, 1995; Lemieux and Bachand, 2009; Calado et al., 2000; Apponi et al., 2010). It has been shown that manipulation of PABPN1 or expanded PABPN1 (expPABPN1) expression levels in muscle cells, i.e. high over-expression or complete gene knockdown, leads to muscle defects including muscle weakness and muscle atrophy, impaired cell growth and apoptosis, and impaired cell fusion (Apponi et al., 2010; Chartier et al., 2006; Davies et al., 2006; Trollet et al., 2010). We have shown that expPABPN1 expression in muscle fibres leads to substantial gene deregulation in OPMD patients and in OPMD model systems (Anvar et al., 2011a; Raz et al., 2011; Trollet et al., 2010). In the OPMD mouse model, by performing an integrative analysis empowered by a number of computational and data-mining methods, we reported muscle atrophy as the major contributor to muscle weakness. This was evident from both the reduction of muscle mass and loss of contractile force due to increased fibrosis, mitochondrial defects, oxidative stress, and deregulation of the ubiquitin-proteasome system (see Chapter **one**). In this mouse model, these observations were mainly restricted to the glycolytic fibres. Despite some pathological similarities between OPMD patients and the mouse model, muscle atrophy is rare among OPMD patients. It is possible that the severe atrophy in glycolytic fibres of OPMD mice is the result of the high overexpression of expPABPN1 rather than the mutation itself. To correct for potential artefacts, an integrative approach was designed in which microarray datasets from three different organisms were combined to gain insight in the common molecular pathways that underlie OPMD. The result of such an extensive strategy, presented in Chapter **two**, was the identification of the ubiquitin-proteasome system (UPS) as the most prominently deregulated molecular pathway in OPMD model systems and patients (Anvar et al., 2011a). Transcriptome studies in non-muscle cells expressing expPABPN1 did not reveal prominent deregulation of UPS genes (Corbeil-Girard et al., 2005). This suggests that the effect of expPABPN1 on UPS deregulation is specific to muscle cells or to post-mitotic cells. Deregulation of the UPS has also been reported in myotonic dystrophy type 1 (Vignaud et al., 2010) and in muscle atrophy in mice (Cao et al., 2005; Bodine et al., 2001; Sandri, 2008). In addition, altered UPS activity has been associated with muscle ageing (Combaret et al., 2009; Lee et al., 1999). Together, these studies suggest that muscle cell function is tightly regulated by the UPS.

The UPS is the main regulator of protein homeostasis (also referred to as proteostasis) and is involved in a wide spectrum of human diseases including cancer, neurodegenerative disorders and diabetes (Hoeller and Dikic, 2009; Liu et al., 2000; Combaret et al., 2009; Taillandier et al., 2004; Ciechanover and Brundin, 2003). To maintain protein homeostasis, it is essential to uphold balance between activities of protein quality-control machineries, the UPS and autophagy-lysosome (Powers et al., 2009). These machineries can adequately respond to damaged proteins and organelles through adjustment of the level of chaperones and proteases (Goldberg, 2003; Meusser et al., 2005; Guisbert et al., 2008; Morimoto, 2008; Ron and Walter, 2007). However, progressive exhaustion of these quality-control systems, owing to ageing (Hipkiss, 2006; Wang et al., 2009; Munch and Bertolotti, 2010; Ben-Zvi et al., 2009) or genetic mutation (Olzmann et al.,

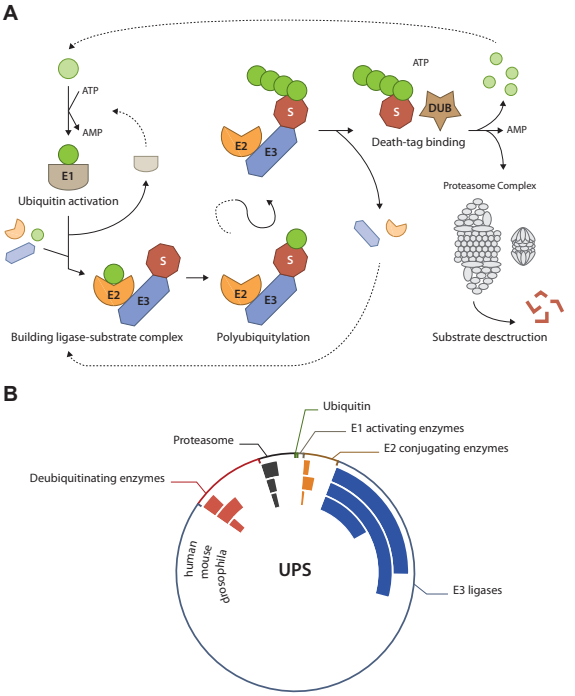

Figure 1 – Schematic overview of the ubiquitin-proteasome system. A) Protein degradation through the ubiquitin-proteasome system involves several steps. Firstly, the ubiquitin (Ub) is being activated by ubiquitin-activating enzyme (E1). Next, ubiquitin is delivered to ubiquitin-conjugation enzyme (E2) for formation of the E2-Ub, ubiquitin ligase (E3) and substrate complex. Consequently, ubiquitins are being transferred to the substrate in order to tag the substrate with the polyubiquitin chain. In the fourth step, E3 releases the polyubiquitylated substrate. The proteasome recognises the polyubiquitin chain as a degradation signal. Therefore, substrate is deubiquitinated and destroyed by the proteasome in ATP-manner. **B)** Within the ubiquitin-proteasome system, E2, E3, deubiquitinating enzymes and proteasome show significant deregulation. Pie charts illustrate the relative distribution of the deregulated genes widespread throughout different components of the ubiquitin-proteasome system across species. A fraction of deregulated genes within individual species are shown in dark colours.

2007), would lead to accumulation of altered proteins as the accumulation of misfolded proteins surpasses the system's capacity (Tyedmers et al., 2010). It has been suggested that the aggregation of proteins can spread to other proteins of mainly the same type (Gidalevitz et al., 2006; Rajan et al., 2001; Ben-Zvi and Goloubinoff, 2002) which could explain the age-dependent and progressive nature of protein aggregation phenomena. Excessive aggregation of proteins could lead to a progressive decline in the level of soluble proteins available in cells. This may result in reduced level of functional protein and, consequently, lead to pathophysiological abnormalities. Thus, understanding the molecular processes regulating cellular homeostasis may unravel mechanistic insights in pathological aspects of various protein aggregation and late-onset diseases.

The UPS involves an enzymatic cascade of ubiquitination and degradation steps. From the six UPS components, only E3-ligases, deubiquitinating enzymes and the proteasome were found to be consistently and prominently deregulated in OPMD model systems and patients (**Figure 1**). Particularly, E3-ligases are fundamental to the specificity of this system and are classified into the RING finger, HECT, and U-Box domains (Deshaies and Joazeiro, 2009). Moreover, many of the E3-ligases play an important role in maintaining genomic integrity (Lipkowitz and Weissman, 2011). Therefore, it is essential to understand what type of E3-ligases are involved in forming the E2-E3 complex to specify the fate of proteins involved in protein aggregation disorders. In OPMD, we found that a subset of the deregulated E3-ligases co-localize with the aggregates of mutant PABPN1. Moreover, their RNA expression profiles correlate with their sequential entrapment in intranuclear inclusion (INI) (Anvar et al., 2011a). It would be essential to look for possible E3-ligases that differentially bind and/or regulate wild-type and mutant PABPN1 in OPMD. This is important since differential regulation of PABPN1 may in part explain the enhanced level of PABPN1 aggregates and reduced level of soluble proteins in OPMD patients. Intriguingly, E3-ligases are also recognised as potential drug targets (Nalepa et al., 2006; Xu and Jaffrey, 2011).

Hence, focused profiling efforts can lead to the identification of ubiquitination events that are regulated by potential therapeutic compounds (Kim et al., 2011; Emanuele et al., 2011).

The process of ubiquitination and degradation through UPS machinery is a modifiable process that can be tuned by the manipulation of specific deubiquitinating enzymes or proteasome activity. This possibility further provides opportunities for therapeutic interventions (Lipkowitz and Weissman, 2011; Crawford et al., 2011). Relevant to OPMD, proteasome activity is reduced during muscle aging (Combaret et al., 2009; Lee et al., 1999; Ferrington et al., 2005) and perhaps, consequently, leads to accumulation of altered proteins. Concordantly, the expression of many aggregation-prone proteins was found to be deregulated in OPMD as well as in other protein aggregation disorders (Anvar et al., 2011a; Corbeil-Girard et al., 2005). Our analysis revealed that the core subunit of the proteasome was consistently down-regulated in OPMD (Anvar et al., 2011a). Additionally, we have shown that expPABPN1 expression in myotubes leads to down-regulation of proteasome-encoding genes and affects the accumulation of expPABPN1 protein (Raz et al., 2011). In turn, manipulation of proteasome activity also affects the accumulation and aggregation of expPABPN1 (Anvar et al., 2011a; Raz et al., 2011). In spite of this prominent link between proteasome activity, expPABPN1 accumulation and INI formation, this process is not specific to muscle cells (Abu-Baker et al., 2003). Since the onset of OPMD coincides with proteasomal down-regulation in ageing muscle, it is possible that the decline in proteasome activity during muscle aging triggers or accelerates expPABPN1 accumulation. Subsequently, in OPMD, aggregation of mutant PABPN1 leads to extensive proteasome down-regulation and entrapment of proteasomal proteins in INIs. This feed forward model along with the onset of skeletal muscle ageing could explain the muscle-specific and INI formation in OPMD (**Figure 2**). Notably, decrease in skeletal muscle performance, as measured by muscle strength, strongly correlates with chronological ageing (Beenakker et al., 2010). Loss of muscle function during ageing is regulated by numerous genetic and environmental factors (Roth et al., 2002) which may explain the differences in muscle performance among individuals (Kostek and Delmonico, 2011). Ageing associated physiological changes can be accompanied by an increased susceptibility to degenerative disorders (Kirkwood and Austad, 2000). Although in most tissues ageing is marked by a progressive decline of cellular functions starting at mid-life (Kirkwood, 2005; Lexell et al., 1988; Lindle et al., 1997; Sahin and Depinho, 2010), the rate of functional changes is tissue-specific (Kirkwood and Austad, 2000).

We found substantial similarities in transcriptional changes between muscle ageing and OPMD. The most striking finding, based on the analysis of expression profiles, was the significant decline in *PABPN1* expression during the first half of the fifth decade. Since changes in skeletal muscle performance commence at the fifth decade (Lindle et al., 1997; Roth et al., 2002) our results suggest a correlation between *PABPN1* expression and the onset of muscle ageing. Moreover, among controls, *PABPN1* expression in females was significantly lower than in males. This observation is in agreement with previous studies indicating that ageing-associated changes in muscle strength are more pronounced in females (Kent-Braun et al., 2002; Roth et al., 2002). Concordantly, the OPMD prevalence of the Uruguayan population is estimated to be higher in females (Medici et al., 1997). Thus, the *PABPN1* expression profile could additionally mark gender-associated decline in muscle performance. Together, the progressive decline in *PABPN1* expression during muscle ageing and the accelerated reduction of its expression in OPMD indicate a strong correlation with muscle weakness. The early mid-life onset of *PABPN1* down-regulation (as compared to that of frontal cortex brain tissues (Lu et al., 2004) with the onset of 85 years; and *Rectus Abdominis* (Zahn et al., 2006) tissues with unchanged expression) suggests temporal-spatial specific-

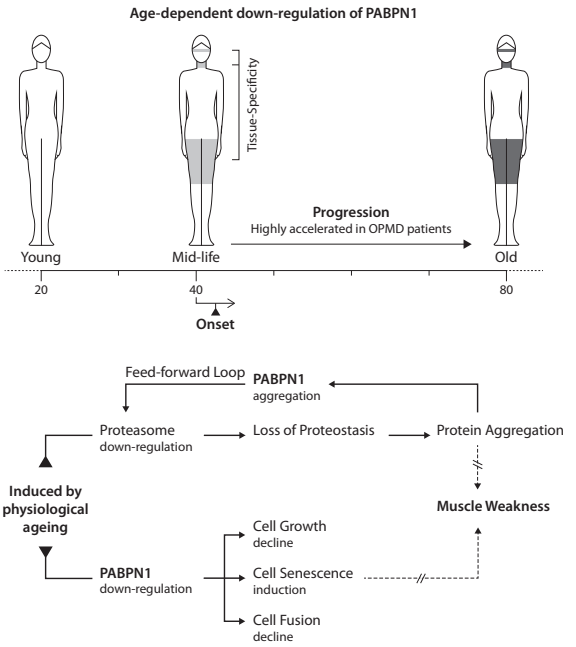

Figure 2 – A model for molecular mechanisms involved in OPMD pathology. In muscles, age-associated proteasome down-regulation triggers expPABPN1 protein accumulation. Subsequently, elevated expPABPN1 aggregation leads to proteasome deregulation during disease onset. This feed forward loop and the onset of muscle ageing leads to loss of proteostasis and INI formation. As part of ageing-related transcriptional changes, there is a significant reduction in the expression of PABPN1. This age-associated decline is accelerated in OPMD patients. In cell cultures, reduced expression of PABPN1 during ageing of skeletal muscles leads to progressive cell senescence and defects in cell fusion and growth. The effect on the expression of muscle contraction genes highly depends on the level of PABPN1 expression. The decline in PABPN1 expression may partially explain the progressive decline in muscle performance during ageing and accelerated muscle weakness in OPMD patients.

ity. However, some reports have indicated mental retardation, cognitive impairment, spinal cord involvement, and dementia in some OPMD patients (Millefiorini and Filippini, 1967; Sarkar et al., 1995; Blumen et al., 2009; Linoli et al., 1991; Mizoi et al., 2011; Dubbioso et al., 2011). Thus, it would be interesting to assess PABPN1 expression in respect to the central nervous system.

Age-dependent progressive decline of PABPN1 expression and loss of muscle function suggests that PABPN1 may play a role in ageing of skeletal muscles (see Chapter **three**). PABPN1 expression in OPMD patients is only 30% of that found in young healthy controls. In immortalized human myoblast cultures, this expression level leads to progressive cellular defects including reduced cell growth and fusion and induced cell senescence. Heterochromatic foci (HF), the hallmark of cellular senescence (Spector and Gasser, 2003), could be observed in cells with 70% PABPN1 down-regulation. Notably, PABPN1 expression was undetectable in nuclei with HF. We suggest that the effect of PABPN1 down-regulation on cellular senescence is more pronounced in non-mitotic cells as they exhibit a three-fold higher amount of cells with HF. Myotube cultures from OPMD muscles also show premature senescence and reduced cell fusion (Perie et al., 2006). Relevant to reduced muscle performance, the expression of muscle contraction genes highly depend on PABPN1 expression level. Recently, we showed that increased PABPN1 protein accumulation in muscle cells results in a reduced amount of the soluble and functional protein (Raz et al., 2011). Since PABPN1 regulates mRNA stability it is expected that decline in functional PABPN1 would have a broad effect on cellular functions as demonstrated here and by Apponi et al. (Apponi et al., 2010). Together, for the first time, our data indicates the progressive response of muscle cell function to the level of PABPN1 in a spatial-temporal manner, highlighting PABPN1 role as a key regulator of muscle ageing.

Ageing cells exhibit distinctive features ranging from the accumulation of damaged macromolecules to changes in nuclear architecture (Campisi and Vijg, 2009; Oberdoerffer and Sinclair,

2007). In particular, it has been suggested that ageing and age-related disorders are strongly associated with mechanisms that control chromatin structure through DNA methylation, RNA interference, histone variants, and post-translational modifications (Oberdoerffer and Sinclair, 2007; Campisi and Vijg, 2009; Rakyan et al., 2010; Teschendorff et al., 2010; Esteller, 2007; Pogribny et al., 2006; Tryndyak et al., 2006; Ronn et al., 2008; Tohgi et al., 1999; Martin, 2009; Rando and Chang, 2012). Nuclear chromatin is associated with processes that mediate DNA replication and transcription (Trinkle-Mulcahy and Lamond, 2007). Markedly, gene transcription is strongly modulated by its relative position within the nucleus (Sexton et al., 2007). This suggests that disruption of the positioning of the chromatin at the nuclear envelop can affect the regulation of gene expression (Akhtar and Gasser, 2007). Furthermore, DNA and chromatin modifications are recognized as both responsive and effectors of the ageing process (Martin, 2009; Rando and Chang, 2012). Therefore, the spatial distribution of genes across the nuclear envelop can significantly contribute to the transcriptional control. Aged cells, in particular, show several changes on their chromatin and nuclear envelop structure that contribute to the lineage and tissue-specific gene expression (Krishnamurthy et al., 2004; Rando and Chang, 2012). It will be interesting to investigate the possibility in which epigenetic changes play a role in mechanisms that underlie the onset and progression of OPMD. Identification of possible epigenetic factors that may be functionally associated with the OPMD phenotype can provide insights on the relationship between the genome and environment. This would potentially lead to a better understanding and characterization of the severity of symptoms in OPMD patients.

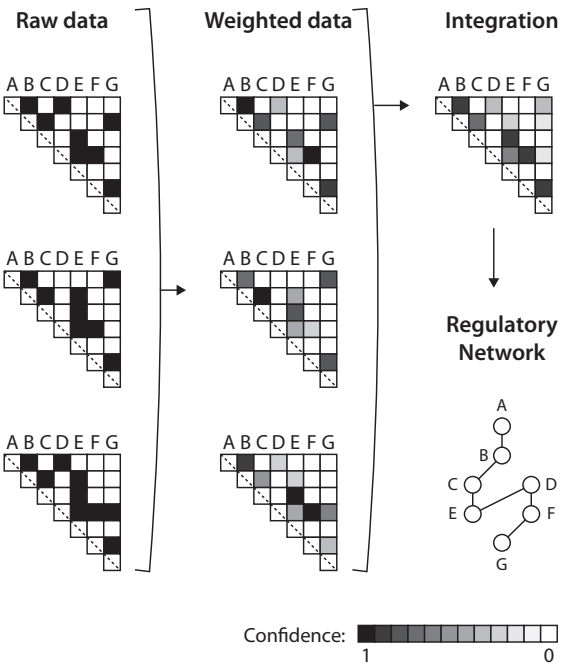

Figure 3 – Probabilistic network integration. Datasets from multiple sets of independent experiments on different species are individually tested and optimised for their association with a given phenotype. Various statistical approaches can be used to infer confidence weight for any given intraspecies regulatory relationship. These weights can then be used to integrate network structures across species. Graphical networks can be derived from the final weighted matrix after applying a confidence threshold.

PABPN1 is involved in pre-mRNA polyadenylation, where it stimulates poly(A) polymerase and regulates poly(A) tail length and RNA stability (Lemay et al., 2010; Kuhn et al., 2009). It is now widely accepted that alternative processing of pre-mRNA can result in structural variation and differing function of encoded proteins (Moore and Silver, 2008; Birzele et al., 2008), as well as regulation of gene expression. Elongation or shortening of the 3' un-translated region (UTR), as a consequence of alternative polyadenylation, can lead to changes in binding of miRNAs and, therefore, differential regulation of mRNAs. Considering the role of PABPN1 in regulating the poly(A) tail and initial indications regarding a widespread discordance between deregulated transcripts of genes, it is crucial to pursue such investigation using next-generation sequencing. Moreover, mechanisms that regulate the 3' UTR are controlled in a tissue-specific manner. There-

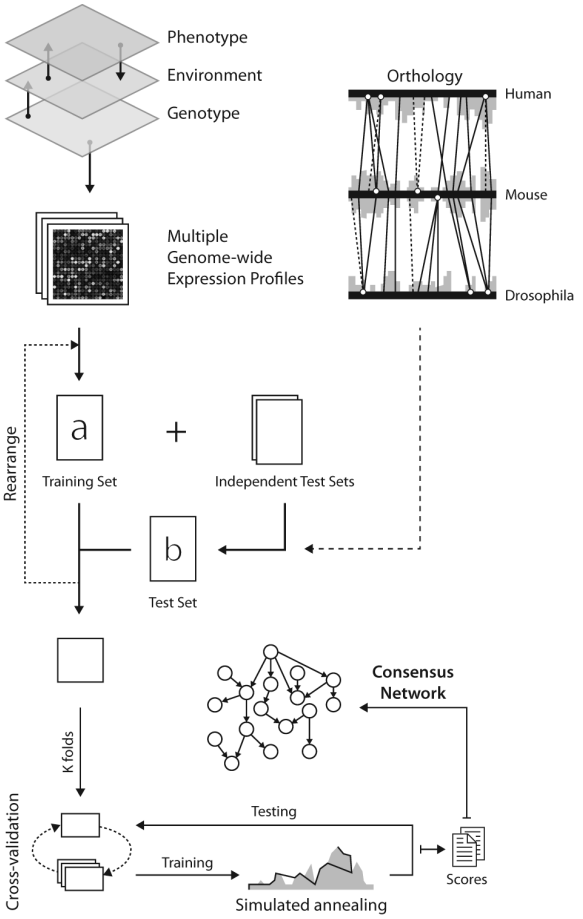

Figure 4 – Schematic overview of the Dandelion algorithm for disease network analysis. The Dandelion algorithm involves three recurring stages of training and an independent testing regime with the use of multiple datasets derived from different species. In the first step, disease modules are defined based on prior knowledge. The next step involves reiterative selection of one species for which the gene regulatory network is constructed while others are left aside for independent testing and validation of the learnt disease networks. For the construction of an intraspecies disease network, the dataset is divided into k-folds, using cross-validation. Subsequent, regulatory relationships between gene transcripts are learnt using a Bayesian network methodology based upon simulated annealing optimization of the network Bayes Information Criterion (BIC) score. After applying confidence thresholds on relationships between genes, the disease network is translated to the expected interspecies disease network. This is achieved by the use of the cross-validation and network optimization procedure. The algorithm searches through the relationships found in the training dataset to find the best fit for the interspecies representation of the disease network. These networks are then integrated by removing all the links with low confidence score across species.

fore, it is essential to pinpoint how ageing-associated decline of soluble PABPN1 and induced aggregation of the mutant PABPN1 may lead to tissue-specific changes in poly(A) site usage. In addition, diversification of RNA, and consequently protein function and structure, is regulated through processes of which alternative splicing plays a central role. In particular, skeletal muscle is reported as one of the tissues with the highest rate of alternative splicing (Pan et al., 2008; Wang et al., 2008; Castle et al., 2008). It is not surprising that genetic mutations may lead to deregulation of this process and consequently cause a widespread transcriptional changes (Cooper et al., 2009; Wang and Cooper, 2007; Tazi et al., 2009). Thus, it is important to pursue the possibility that alternative splicing is differentially regulated in OPMD patients and model systems.

The work, presented in this thesis, strongly highlights the fascinating nature and value of interdisciplinary studies. We have shown that the concept of a universality of biological processes in the light of evolutionary mechanisms and common functional processes can lead to novel discoveries. Engaging in the study of a variety of organisms or biological behaviours, looking for shared molecular features in a rare disease such as OPMD, enabled us to uncover insights on a broader spectrum of conditions and phenomena such as ageing of skeletal muscles and protein aggregation disorders.

The inner workings of complex biological networks

Functional interdependencies and the modular nature of cell's molecular components imply the indispensable role of network-based approaches to human diseases. It is widely established that these dependencies revealed by regulatory networks can provide valuable information regarding underlying biological processes (Avery and Wasserman, 1992; Tong et al., 2004; Costanzo et al., 2010; St Onge et al., 2007; Schuldiner et al., 2008; Collins et al., 2007; Segre et al., 2005; Drees et al., 2005; Guarente, 1993; Hartman et al., 2001; Jonikas et al., 2009). Despite the generation of vast quantities of data by high-throughput technologies, biological data are usually sparse, noisy and ambiguous, limited in number of samples, and high-dimensional. Thus, integration of data and genomic information from human and various model systems can ultimately provide a better indication of common molecular mechanisms that underlie a given phenotype. However, the presence of noise and technical artefacts specific to model systems usually leads to limited overlap between results obtained in cross-species comparison (Lu et al., 2009; Zhou and Gibson, 2004; Oliva et al., 2005; Blake et al., 2003; Jelier et al., 2008). Additionally, integrative approaches are far from trivial and are complicated due to our limited knowledge of true protein orthologues, transcript variants coding for proteins with similar function, and evolutionary conservation of biological processes. These bottlenecks further require fine tuning and optimization of the integration strategy. Another aspect of complexity arises from the generation of large-scale networks (having thousands of nodes and millions of possible interactions), owing to limited computational power and intelligent algorithms for scalability and reducing dimensionality (Venkatesan et al., 2009; Barabasi et al., 2011). Markedly, such stochastic systems require a probabilistic approach at the core for modelling regulatory networks.

We first established, in Chapter **four**, a way in which gene networks that are highly informative for determining "muscle differentiation" can be robustly identified from multiple independent datasets with increasing level of complexity and stochasticity (Anvar et al., 2010). We showed that the proper use of a modelling strategy in combination with multiple datasets leads to the construction of gene networks that can explain the myogenesis-related genes significantly better than those that have less involvement in myogenesis. This approach resulted in networks that were consistently more parsimonious to myogenesis-related genes. Moreover, these models provide the robust prediction of biological outcome and expression profiles. Establishing a strategy which can accommodate the integration of multiple datasets enables the possibility of overcoming the limitations of cross-species integrative studies. Such exploitation would lead to more robust regulatory mechanisms to be identified and predictions to be made across various platforms and organisms (**Figure 3** and **Figure 4**). In Chapter **five**, we showed that the integration and analysis of microarray datasets from various species increase the robustness of the constructed networks and the predictive accuracy of the disease state (Anvar et al., 2011b). We also demonstrated that the interspecies translation of these networks helps to avoid overfitting. In addition, this approach provides a state-of-the-art model-driven selection of transcript isoforms that are most likely to be coding for orthologous proteins. Notably, another fascinating application of this strategy would be the identification of alternative splicing events and their regulators (Zhang et al., 2010). These powerful features are essential for understanding the phenotypic implications of such strong relationships as part of evaluating the conservation and dynamics of interspecies disease networks. Moreover, the high level of specificity and sensitivity of these models enables the prioritization of candidate regulators of the disease molecular mechanisms to be studied in follow-up validation experiments. In particular, it is crucial to carry out additional experiments to investigate the tissue-specificity of the network (Reverter et al., 2008; Lage et al., 2010) and the functional relevance of encoded proteins dysregulation to the disease pathology. This can also

be achieved by re-constructing tissue- or cell-specific sub-networks from the model by integrating a variety of tissue-specific data sources (Jerby et al., 2010; Kirouac et al., 2010).

Our approach for Bayesian modelling of datasets on a similar phenotype from different model systems and patients is unique. Several approaches have been described to avoid overfitting and increase the robustness of Bayesian networks. For example, informative priors derived from protein-protein interaction (PPI) data or from the literature have been used to generate more stable and biologically meaningful networks (Segal et al., 2003; Pe'er et al., 2002; Steele et al., 2009; Jansen et al., 2003). While these methods obviously bias the results towards well-known regulatory interactions and are less likely to detect novel relationships (Sprinzak et al., 2003; Joyce and Palsson, 2006), they may ultimately be combined with our modelling approach to obtain regulatory networks with a more straightforward biological interpretation.

Our method was applied to an *a priori* defined gene module coding for a well-known biological structure, the proteasome. Several studies in *S. cerevisiae* (Zhang et al., 2005; Tanay et al., 2004; Luscombe et al., 2004; Han et al., 2004) have demonstrated the value of an integrative modelling approach providing modularized interaction networks without prior assumptions. Zhang et al. (Zhang et al., 2005), for instance, took an approach in which they integrated a number of different available data sources, from PPIs to sequence homology and gene co-expression, while Tanay et al. (Tanay et al., 2004) and others (Luscombe et al., 2004; Han et al., 2004) expanded on the statistical analysis of network properties and identified modules within the network structure. The performance of

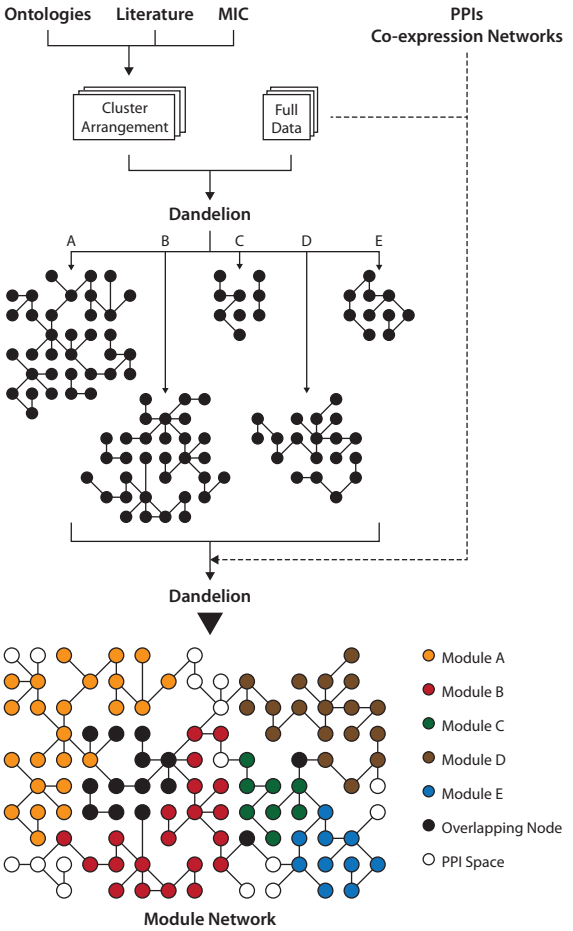

Figure 5 – Schematic illustration of Dandelion module networks. In a bottom-up approach, gene modules are curated on the basis of their literature-aided and cross-species association or according to predefined ontologies. The algorithm involves three recurring phases of training and an independent testing regime with the use of multiple datasets from different platforms, experiments, or organisms. First, consensus networks are constructed for individual modules using our previously described Dandelion algorithm (Chapter **five**). These sub-networks are then overlaid based on common nodes and relationships within different network structures. Finally, using protein-protein interaction databases or associations in co-expression networks, the Dandelion algorithm attempts to assemble and optimise the full module network by adding relationships and nodes to interlink sub-networks. Additionally, the Dandelion algorithm would allow for novel nodes and relationships to be added to the global module network structure. The growth of module networks is constrained on the overall improvement of the network performance.

these models depends on the availability of high quantities of samples and may be prone to overfitting due to the presence of noise and other model-specific artefacts. Therefore, a combination of their approach with our interspecies translation may enable the discovery of larger gene regulatory networks with multiple gene modules and connections between them.

Understanding the dynamics of network structure is essential for determining causal interdependencies as well as characterization of network modularity and gene spatial properties. In model organisms, it has been shown that hub proteins are tend to be encoded by essential genes (Jeong et al., 2001) which are highly conserved (Fraser et al., 2002; Eisenberg and Levanon, 2003; Saeed and Deane, 2006). Identification of essential genes is important for discovery of sub-networks that are associated with a disease phenotype, owing to the disease-related genes being located in the network-based vicinity of the hub nodes (Goh et al., 2007; Feldman et al., 2008). The importance of nodes to the network can be estimated using the 'betweenness centrality' measure (Yu et al., 2007) which gives some additional insights on topology, information flow, and the stability of a network (Han, 2008). Topological properties of disease networks reveal clouds of densely interconnected nodes that can be used for gene module prediction (Girvan and Newman, 2002; Palla et al., 2005; Ahn et al., 2010; Enright et al., 2002). In addition to network topology, functional characterisation of sub-networks can improve in describing mechanisms that give rise to a specific phenotype.

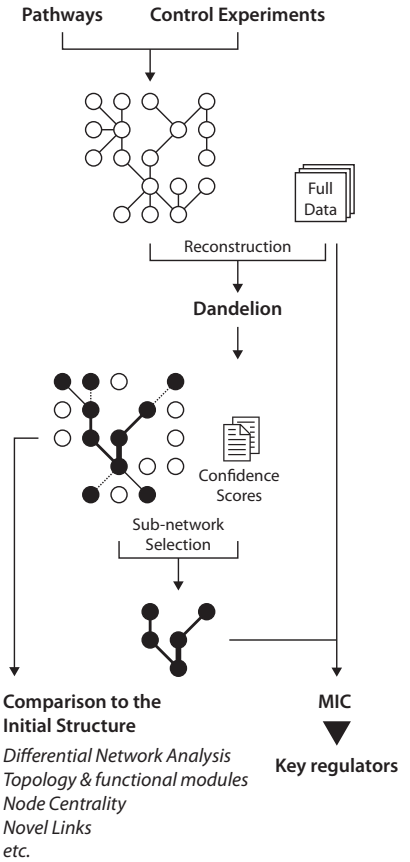

Figure 6 – A model for network reconstruction and evaluation. Known networks, produced based on control experiments or molecular pathways, can be reconstructed using the Dandelion algorithm. Reconstructed networks differ in respect to nodes being incorporated (depicted in black) within the network structure. Moreover, the comparative analysis can unravel changes in the dynamics of such regulatory networks under different phenotype or experimental settings. Relationships with the strongest weights (depicted by the line thickness), on the basis of their confidence score, can be evaluated using the maximal information coefficient (MIC) measure. These strategies could uncover key regulators of a given pathway under specific phenotype or experimental setting.

Here, I discuss a strategy to tackle some challenges in bridging the gap between multi-layers of biological data. In a study presented in Chapter **five**, we developed a novel algorithm for constructing interspecies disease networks that provide an assumption-free and model-driven selection of the most important transcript isoforms across species (Anvar et al., 2011b). We achieved this by use of prior knowledge on pathways that are disease-associated. This was owing to the fact that, on a genome-wide scale, searching the space of possible networks via single-arc changes is not realistic and computationally expensive. One of the possible strategies for reducing the high dimensionality is the use of statistical algorithms such as ridge and LASSO regression (Friedman et al., 2008; Tibshirani, 1996). These algorithms apply a penalty for complex models that may be tuned by cross-validation. However, this would mean that the same dataset is

used, in two steps, for generating the network space and construction of disease network which would lead to overfitting and other biases. Another alternative is based on treating functional modules as blocks of interconnected nodes which can be assembled together by the use of overlaying nodes and relationships. In conjunction with assembling overlaying nodes and modules, additional links are added to the network for global optimization of the inter-module relationships (**Figure 5**). This can be reached by simple hill-climbing, greedy algorithms or more sophisticated simulated annealing and MCMC (Markov Chain Monte Carlo) searching methods. Within the optimization step, evidences from PPI networks can be used for confidence assessment. In addition to the utility of PPI networks, reconstruction of known functional pathways, or those produced by alternative models on control datasets, can be combined with allowing for novel relationships (Battle et al., 2010) (**Figure 6**). This strategy potentially can help automating the process of optimization and confidence assessment. Moreover, the maximal information coefficient (Reshef et al., 2011) can be integrated to assess the functional association for relationships in the vicinity of the essential nodes.

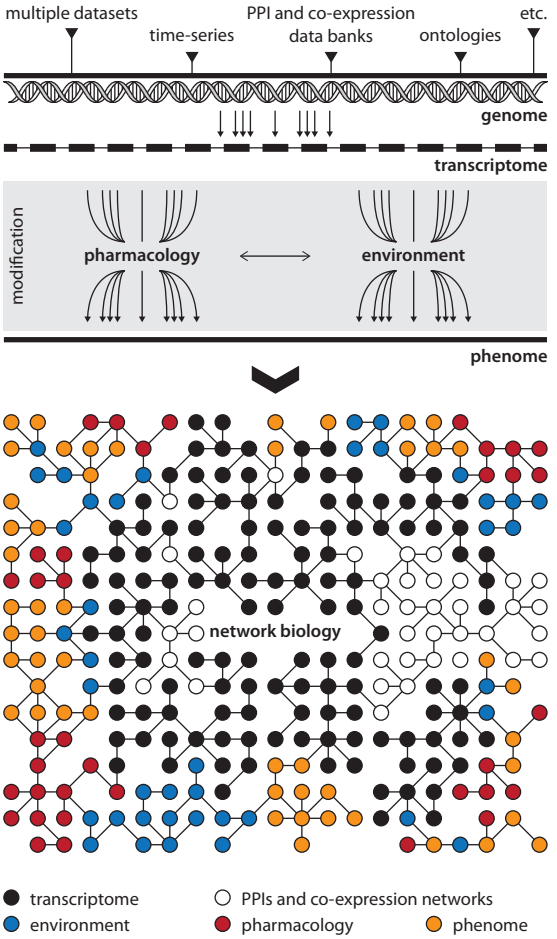

Figure 7 – Network medicine, linking across multi-layers of biological data. Datasets from different organisms or platforms can be combined for enhanced identification of interspecies (inter-platform) networks. Time-series datasets can provide information on the dynamics of biological networks while protein-protein interaction and co-expression networks can be used for optimization and scaling. Networks constructed on transcriptome data are linked to networks related to pharmacology, phenomics, and environment. Genomic information, reflected in transcriptome, are interlinked and translated to diverse sets of phenotypes through environmental factors. Likewise, these multi-layers of densely interconnected regulatory relationships are represented through a framework of pharmacological entities.

Finally, the evolution of these network properties over time would provide a crucial framework for better understanding the causal relationship and dynamics of gene regulatory networks in the context of human diseases. Thus, I believe that robust and unbiased construction and analysis of the interspecies networks for rare or complex human diseases can lead to novel discovery and identification of key regulators. The result of such exploration can ultimately offer potential targets for therapeutic interventions and drug developments. In the last section, I will discuss a few strategies that, in my view, can be pursued to enhance data integration and the ideal utility of network-based approaches on a larger scale. This would consequently provide a disease-oriented global view of genomics, transcriptomics, proteomics, and newly defined field of phenomics. The term

"phenomics" is derived from the word 'phenome' which was first introduced by Michael Soule (Soule, 1967).

From systems biology to personalized medicine

In recent years, the studies of human diseases have changed significantly, owing to advancements in the field of systems biology and high-throughput technologies. It is widely believed that the integration of genomics, transcriptomics, and large-scale phenotyping has major potential for novel discoveries in network biology of multi-layered human disorders (Bilder et al., 2009; Freimer and Sabatti, 2003; Schilling et al., 1999; Searls, 2005). Phenotypic variations are determined by a complex network of genetic and environmental interactions. In the last two decades, significant efforts have been made on genomic and transcriptomic studies. Now we have reached the time to invest special efforts in the field of phenomics which still requires our careful attention. This is due to the lack of standardisation and data available. Although limited phenotyping efforts for obvious disease-related features seems to be sufficient in most cases, extensive and global phenotyping can pave the way for better standardization of phenotypic information and mechanistic understanding of context-oriented genetic and environmental interactions. This would lead to the discovery of novel dependencies between genomics, transcriptomics, and phenomics data. Additionally, this combinatory framework provides an information-rich model that can distinctly characterise the correlation or causal relationships and account for different sources of variation (Houle et al., 2010). The importance of the integrative approaches is evident from experiments carried out in yeast that demonstrate a substantial growth from lethal or disease causing single-gene deletions (34%) to those that occur in conjunction with at least one environmental condition (97%) (Hopkins, 2008). Nevertheless, the design of an integrative strategy needs to be addressed with precision and care as navigating such data is extremely challenging. For instance, one of the basic information losses is that phenotypic data is often treated as a discretised entity whereas the most vital piece of information lays in the relative and continuous changes of phenotypic information, phenomena which is now well-established for other data-types such as transcriptome.

Having established the modular network structure, the next step in exploring the interplay between transcriptomics and phenotypic states of human diseases is to determine the environmental factors through which these networks are regulated in a full range of spatial and temporal scale. Adequate combination of prior knowledge (Ochs, 2010) can further provide confirmatory insights on data-flow and ordering of causal relationships across multi-layers of biological data (**Figure 7**). Yet, the use of prior knowledge-centric approaches needs to be avoided to minimize the biases that can be introduced by such techniques. An intriguing possibility of such methods is that the construction of multifaceted biological networks may provide insights on efficacy and off-target toxicity of drugs in a phenotype-centric and tissue-specific manner, some of which can be determined by the analysis of such network structure (Albert et al., 2000; Kitano, 2007). Likewise, special efforts in modelling the dynamics of metabolic responses in different tissues can provide valuable insights into the effects of drugs and diseases (**Figure 8**). Another intriguing benefit of engaging in metabolomics studies is the possibility of linking different levels of biological organization (genomics, transcriptomics, proteomics, etc.), owing to their differing operational behaviour (Holmes et al., 2008a; Holmes et al., 2008b; Nicholson and Wilson, 2003). An extensive review by Hopkins (Hopkins, 2008) provides valuable information on the usability of biological networks in drug discovery along with a brief outlook on future prospects. While some of these advancements seem farfetched and years in the future, a few preliminary developments can be pursued that provides a new basis for a global infrastructure of network medicine. For in-

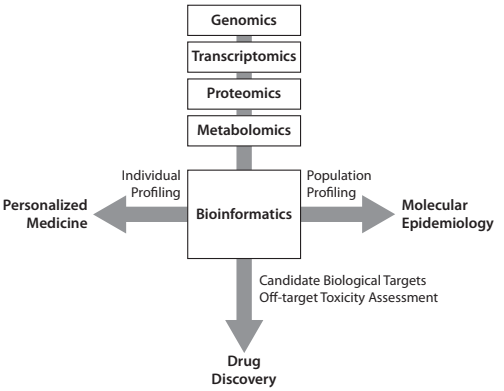

Figure 8 – Applications of multifaceted omics research. Special efforts in combined analysis across multi-layers of biological data provide an infrastructure in which bioinformatics can play a central role. Naturally, the main applications of omics research can be divided into three fields of personalized medicine, drug discovery, and molecular epidemiology. Profiling of individuals can provide an enhanced framework for better therapeutic interventions. The utility of this strategy is to comprehend patients' susceptibility to diseases or alter therapies on the basis of their response to different medicine. Molecular epidemiology studies can be enhanced by looking for common patterns in profiles within a population. This would allow for the identification of biomarkers, susceptibilities of specific populations to diseases, and health screening programmes. Finally, these studies can lead to uncovering new biological targets for drug discovery.

stance, the adoption of methods that deal with dynamics of these networks, in a spatial-temporal manner, can act as a cornerstone for robust integration of pharmaceutical data and chemical interactions. This combinatory strategy provides a valuable framework for drug discovery and personalized therapeutic interventions. Notably, recent approaches for simple characterization of the network topology had made a remarkable contribution in developing strategies for prioritization and combination of drug targets (Gerber et al., 2008; Potapov et al., 2008; Wunderlich and Mirny, 2006). Mining biomedical and biochemical literature in conjunction with ontologies (such as KEGG and GO) are also well-explored to better determine the efficacy of drug development (Yildirim et al., 2007; Ji et al., 2007; Spiro et al., 2008; Gunther et al., 2008). Bayesian approaches can bridge between these different sources of information and provide a global network infrastructure in which transcriptome data, environmental factors, and phenotypic information can come together to provide a predictive and model-driven framework for assessing the clusters of chemical networks and pharmacology data (**Figure 7**). To conclude, the context- and case-specific identification of the optimal point of interaction between molecules for drug discovery is the future of systems biology applications in the field of personalized medicine. In order to achieve this ambition, novel and integrative advancements are needed to better understand the global organisation of networks in the study of human genetic disorders.

Reference List

Abu-Baker,A., Messaed,C., Laganiere,J., Gaspar,C., Brais,B., and Rouleau,G.A. (2003). Involvement of the ubiquitin-proteasome pathway and molecular chaperones in oculopharyngeal muscular dystrophy. Hum. Mol. Genet *12*, 2609-2623.

Ahn,Y.Y., Bagrow,J.P., and Lehmann,S. (2010). Link communities reveal multiscale complexity in networks. Nature *466*, 761-764.

Akhtar,A. and Gasser,S.M. (2007). The nuclear envelope and transcriptional control. Nat Rev Genet. *8*, 507-517.

Albert,R., Jeong,H., and Barabasi,A.L. (2000). Error and attack tolerance of complex networks. Nature *406*, 378-382.

Anvar,S.Y., 't Hoen,P.A., and Tucker,A. (2010). The identification of informative genes from multiple datasets with increasing complexity. BMC. Bioinformatics. *11*, 32.

Anvar,S.Y., 't Hoen,P.A., Venema,A., van der Sluijs,B., van,E.B., Snoeck,M., Vissing,J., Trollet,C., Dickson,G., Chartier,A., Simonelig,M., van Ommen,G.J., van der Maarel,S.M., and Raz,V. (2011a). Deregulation of the ubiquitin-proteasome system is the predominant molecular pathology in OPMD animal models and patients. Skelet. Muscle *1*, 15.

Anvar,S.Y., Tucker,A., Vinciotti,V., Venema,A., van Ommen,G.J., van der Maarel,S.M., Raz,V., and 't Hoen,P.A. (2011b). Interspecies translation of disease networks increases robustness and predictive accuracy. PLoS. Comput. Biol. *7*, e1002258.

Apponi,L.H., Leung,S.W., Williams,K.R., Valentini,S.R., Corbett,A.H., and Pavlath,G.K. (2010). Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. Hum. Mol. Genet. *19*, 1058-1065.

Avery,L. and Wasserman,S. (1992). Ordering gene function: the interpretation of epistasis in regulatory hierarchies. Trends Genet. *8*, 312-316.

Barabasi,A.L., Gulbahce,N., and Loscalzo,J. (2011). Network medicine: a network-based approach to human disease. Nat Rev Genet. *12*, 56-68.

Battle,A., Jonikas,M.C., Walter,P., Weissman,J.S., and Koller,D. (2010). Automated identification of pathways from quantitative genetic interaction data. Mol. Syst. Biol. *6*, 379.

Beenakker,K.G., Ling,C.H., Meskers,C.G., de Craen,A.J., Stijnen,T., Westendorp,R.G., and Maier,A.B. (2010). Patterns of muscle strength loss with age in the general population and patients with a chronic inflammatory state. Ageing Res. Rev *9*, 431-436.

Ben-Zvi,A., Miller,E.A., and Morimoto,R.I. (2009). Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. Proc. Natl. Acad. Sci. U. S. A *106*, 14914-14919.

Ben-Zvi,A.P. and Goloubinoff,P. (2002). Proteinaceous infectious behavior in non-pathogenic proteins is controlled by molecular chaperones. J. Biol. Chem. *277*, 49422-49427.

Bilder,R.M., Sabb,F.W., Cannon,T.D., London,E.D., Jentsch,J.D., Parker,D.S., Poldrack,R.A., Evans,C., and Freimer,N.B. (2009). Phenomics: the systematic study of phenotypes on a genome-wide scale. Neuroscience *164*, 30-42.

Birzele,F., Csaba,G., and Zimmer,R. (2008). Alternative splicing and protein structure evolution. Nucleic Acids Res. *36*, 550-558.

Blake,W.J., KAErn,M., Cantor,C.R., and Collins,J.J. (2003). Noise in eukaryotic gene expression. Nature *422*, 633-637.

Blumen,S.C., Bouchard,J.P., Brais,B., Carasso,R.L., Paleacu,D., Drory,V.E., Chantal,S., Blumen,N., and Braverman,I. (2009). Cognitive impairment and reduced life span of oculopharyngeal muscular dystrophy homozygotes. Neurology *73*, 596-601.

Bodine,S.C., Latres,E., Baumhueter,S., Lai,V.K., Nunez,L., Clarke,B.A., Poueymirou,W.T., Panaro,F.J., Na,E., Dharmarajan,K., Pan,Z.Q., Valenzuela,D.M., DeChiara,T.M., Stitt,T.N., Yancopoulos,G.D., and Glass,D.J. (2001). Identification of ubiquitin ligases required for skeletal muscle atrophy. Science *294*, 1704-1708.

Calado,A., Kutay,U., Kuhn,U., Wahle,E., and Carmo-Fonseca,M. (2000). Deciphering the cellular pathway for transport of poly(A)-binding protein II. RNA. *6*, 245-256.

Campisi,J. and Vijg,J. (2009). Does damage to DNA and other macromolecules play a role in aging? If so, how? J. Gerontol. A Biol. Sci. Med. Sci. *64*, 175-178.

Cao,P.R., Kim,H.J., and Lecker,S.H. (2005). Ubiquitin-protein ligases in muscle wasting. Int. J. Biochem. Cell Biol. *37*, 2088-2097.

Castle,J.C., Zhang,C., Shah,J.K., Kulkarni,A.V., Kalsotra,A., Cooper,T.A., and Johnson,J.M. (2008). Expression of 24,426 human alternative splicing events and predicted cis regulation in 48 tissues and cell lines. Nat Genet. *40*, 1416-1425.

Chartier,A., Benoit,B., and Simonelig,M. (2006). A Drosophila model of oculopharyngeal muscular dystrophy reveals intrinsic toxicity of PABPN1. EMBO J *25*, 2253-2262.

Ciechanover,A. and Brundin,P. (2003). The ubiquitin proteasome system in neurodegenerative diseases: sometimes the chicken, sometimes the egg. Neuron *40*, 427-446.

Collins,S.R., Miller,K.M., Maas,N.L., Roguev,A., Fillingham,J., Chu,C.S., Schuldiner,M., Gebbia,M., Recht,J., Shales,M., Ding,H., Xu,H., Han,J., Ingvarsdottir,K., Cheng,B., Andrews,B., Boone,C., Berger,S.L., Hieter,P., Zhang,Z., Brown,G.W., Ingles,C.J., Emili,A., Allis,C.D., Toczyski,D.P., Weissman,J.S., Greenblatt,J.F., and Krogan,N.J. (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature *446*, 806-810.

Combaret,L., Dardevet,D., Bechet,D., Taillandier,D., Mosoni,L., and Attaix,D. (2009). Skeletal muscle proteolysis in aging. Curr. Opin. Clin. Nutr. Metab Care *12*, 37-41.

Cooper,T.A., Wan,L., and Dreyfuss,G. (2009). RNA and disease. Cell *136*, 777-793.

Corbeil-Girard,L.P., Klein,A.F., Sasseville,A.M., Lavoie,H., Dicaire,M.J., Saint-Denis,A., Page,M., Duranceau,A., Codere,F., Bouchard,J.P., Karpati,G., Rouleau,G.A., Massie,B., Langelier,Y., and Brais,B. (2005). PABPN1 overexpression leads to upregulation of genes encoding nuclear proteins that are sequestered in oculopharyngeal muscular dystrophy nuclear inclusions. Neurobiol. Dis. *18*, 551-567.

Costanzo,M., Baryshnikova,A., Bellay,J., Kim,Y., Spear,E.D., Sevier,C.S., Ding,H., Koh,J.L., Toufighi,K., Mostafavi,S., Prinz,J., St Onge,R.P., VanderSluis,B., Makhnevych,T., Vizeacoumar,F.J., Alizadeh,S., Bahr,S., Brost,R.L., Chen,Y., Cokol,M., Deshpande,R., Li,Z., Lin,Z.Y., Liang,W., Marback,M., Paw,J., San Luis,B.J., Shuteriqi,E., Tong,A.H., van,D.N., Wallace,I.M., Whitney,J.A., Weirauch,M.T., Zhong,G., Zhu,H., Houry,W.A., Brudno,M., Ragibizadeh,S., Papp,B., Pal,C., Roth,F.P., Giaever,G., Nislow,C., Troyanskaya,O.G., Bussey,H., Bader,G.D., Gingras,A.C., Morris,Q.D., Kim,P.M., Kaiser,C.A., Myers,C.L., Andrews,B.J., and Boone,C. (2010). The genetic landscape of a cell. Science *327*, 425-431.

Crawford,L.J., Walker,B., and Irvine,A.E. (2011). Proteasome inhibitors in cancer therapy. J. Cell Commun. Signal. *5*, 101-110.

Davies,J.E., Sarkar,S., and Rubinsztein,D.C. (2006). Trehalose reduces aggregate formation and delays pathology in a transgenic mouse model of oculopharyngeal muscular dystrophy. Hum. Mol. Genet *15*, 23-31.

Deshaies,R.J. and Joazeiro,C.A. (2009). RING domain E3 ubiquitin ligases. Annu. Rev Biochem. *78*, 399-434.

Drees,B.L., Thorsson,V., Carter,G.W., Rives,A.W., Raymond,M.Z., Avila-Campillo,I., Shannon,P., and Galitski,T. (2005). Derivation of genetic interaction networks from quantitative phenotype data. Genome Biol. *6*, R38.

Dubbioso,R., Moretta,P., Manganelli,F., Fiorillo,C., Iodice,R., Trojano,L., and Santoro,L. (2011). Executive functions are impaired in heterozygote patients with oculopharyngeal muscular dystrophy. J. Neurol.

Eisenberg,E. and Levanon,E.Y. (2003). Preferential attachment in the protein network evolution. Phys. Rev Lett. *91*, 138701.

Emanuele,M.J., Elia,A.E., Xu,Q., Thoma,C.R., Izhar,L., Leng,Y., Guo,A., Chen,Y.N., Rush,J., Hsu,P.W., Yen,H.C., and Elledge,S.J. (2011). Global identification of modular cullin-RING ligase substrates. Cell *147*, 459-474.

Enright,A.J., Van,D.S., and Ouzounis,C.A. (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. *30*, 1575-1584.

Esteller,M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet. *8*, 286-298.

Feldman,I., Rzhetsky,A., and Vitkup,D. (2008). Network properties of genes harboring inherited disease mutations. Proc. Natl. Acad. Sci. U. S. A *105*, 4323-4328.

Ferrington,D.A., Husom,A.D., and Thompson,L.V. (2005). Altered proteasome structure, function, and oxidation in aged muscle. FASEB J. *19*, 644-646.

Fraser,H.B., Hirsh,A.E., Steinmetz,L.M., Scharfe,C., and Feldman,M.W. (2002). Evolutionary rate in the protein interaction network. Science *296*, 750-752.

Freimer,N. and Sabatti,C. (2003). The human phenome project. Nat Genet. *34*, 15-21.

Friedman,J., Hastie,T., and Tibshirani,R. (2008). Sparse inverse covariance estimation with the graphical lasso. Biostatistics. *9*, 432-441.

Gerber,S., Assmus,H., Bakker,B., and Klipp,E. (2008). Drug-efficacy depends on the inhibitor type and the target position in a metabolic network--a systematic study. J. Theor. Biol. *252*, 442-455.

Gidalevitz,T., Ben-Zvi,A., Ho,K.H., Brignull,H.R., and Morimoto,R.I. (2006). Progressive disruption of cellular protein folding in models of polyglutamine diseases. Science *311*, 1471-1474.

Girvan,M. and Newman,M.E. (2002). Community structure in social and biological networks. Proc. Natl. Acad. Sci. U. S. A *99*, 7821-7826.

Goh,K.I., Cusick,M.E., Valle,D., Childs,B., Vidal,M., and Barabasi,A.L. (2007). The human disease network. Proc. Natl. Acad. Sci. U. S. A *104*, 8685-8690.

Goldberg,A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. Nature *426*, 895-899.

Guarente,L. (1993). Synthetic enhancement in gene interaction: a genetic tool come of age. Trends Genet. *9*, 362-366.

Guisbert,E., Yura,T., Rhodius,V.A., and Gross,C.A. (2008). Convergence of molecular, modeling, and systems approaches for an understanding of the Escherichia coli heat shock response. Microbiol. Mol. Biol. Rev *72*, 545-554.

Gunther,S., Kuhn,M., Dunkel,M., Campillos,M., Senger,C., Petsalaki,E., Ahmed,J., Urdiales,E.G., Gewiess,A., Jensen,L.J., Schneider,R., Skoblo,R., Russell,R.B., Bourne,P.E., Bork,P., and Preissner,R. (2008). SuperTarget and Matador: resources for exploring drug-target relationships. Nucleic Acids Res. *36*, D919-D922.

Han,J.D. (2008). Understanding biological functions through molecular networks. Cell Res. *18*, 224-237.

Han,J.D., Bertin,N., Hao,T., Goldberg,D.S., Berriz,G.F., Zhang,L.V., Dupuy,D., Walhout,A.J., Cusick,M.E., Roth,F.P., and Vidal,M. (2004). Evidence for dynamically organized modularity in the yeast protein-protein interaction network. Nature *430*, 88-93.

Hartman,J.L., Garvik,B., and Hartwell,L. (2001). Principles for the buffering of genetic variation. Science *291*, 1001-1004.

Hipkiss,A.R. (2006). Accumulation of altered proteins and ageing: causes and effects. Exp. Gerontol. *41*, 464-473.

Hoeller,D. and Dikic,I. (2009). Targeting the ubiquitin system in cancer therapy. Nature *458*, 438-444.

Holmes,E., Loo,R.L., Stamler,J., Bictash,M., Yap,I.K., Chan,Q., Ebbels,T., De,I.M., Brown,I.J., Veselkov,K.A., Daviglus,M.L., Kesteloot,H., Ueshima,H., Zhao,L., Nicholson,J.K., and Elliott,P. (2008a). Human metabolic phenotype diversity and its association with diet and blood pressure. Nature *453*, 396-400.

Holmes,E., Wilson,I.D., and Nicholson,J.K. (2008b). Metabolic phenotyping in health and disease. Cell *134*, 714-717.

Hopkins,A.L. (2008). Network pharmacology: the next paradigm in drug discovery. Nat Chem. Biol. *4*, 682-690.

Houle,D., Govindaraju,D.R., and Omholt,S. (2010). Phenomics: the next challenge. Nat Rev Genet. *11*, 855-866.

Ioannidis,J.P. (2005). Why most published research findings are false. PLoS. Med. *2*, e124.

Ioannidis,J.P. and Khoury,M.J. (2011). Improving validation practices in “omics” research. Science *334*, 1230-1232.

Jansen,R., Yu,H., Greenbaum,D., Kluger,Y., Krogan,N.J., Chung,S., Emili,A., Snyder,M., Greenblatt,J.F., and Gerstein,M. (2003). A Bayesian networks approach for predicting protein-protein interactions from genomic data. Science *302*, 449-453.

Jelier,R., ‘t Hoen,P.A., Sterrenburg,E., den Dunnen,J.T., van Ommen,G.J., Kors,J.A., and Mons,B. (2008). Literature-aided meta-analysis of microarray data: a compendium study on muscle development and disease. BMC. Bioinformatics. *9*, 291.

Jeong,H., Mason,S.P., Barabasi,A.L., and Oltvai,Z.N. (2001). Lethality and centrality in protein networks. Nature *411*, 41-42.

Jerby,L., Shlomi,T., and Ruppin,E. (2010). Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. Mol. Syst. Biol. *6*, 401.

Ji,H.F., Kong,D.X., Shen,L., Chen,L.L., Ma,B.G., and Zhang,H.Y. (2007). Distribution patterns of small-molecule ligands in the protein universe and implications for origin of life and drug discovery. Genome Biol. *8*, R176.

Jonikas,M.C., Collins,S.R., Denic,V., Oh,E., Quan,E.M., Schmid,V., Weibezahn,J., Schwappach,B., Walter,P., Weissman,J.S., and Schuldiner,M. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science *323*, 1693-1697.

Joyce,A.R. and Palsson,B.O. (2006). The model organism as a system: integrating ‘omics’ data sets. Nat Rev Mol. Cell Biol. *7*, 198-210.

Kaplan, W. and Laing, R. Priority Medicines for Europe and the World. 2004. Geneva, World Health Organization.

Ref Type: Report

Kent-Braun,J.A., Ng,A.V., Doyle,J.W., and Towse,T.F. (2002). Human skeletal muscle responses vary with age and gender during fatigue due to incremental isometric exercise. J. Appl. Physiol *93*, 1813-1823.

Kim,W., Bennett,E.J., Huttlin,E.L., Guo,A., Li,J., Possemato,A., Sowa,M.E., Rad,R., Rush,J., Comb,M.J., Harper,J.W., and Gygi,S.P. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol. Cell *44*, 325-340.

Kirkwood,T.B. (2005). Understanding the odd science of aging. Cell *120*, 437-447.

Kirkwood,T.B. and Austad,S.N. (2000). Why do we age? Nature *408*, 233-238.

Kirouac,D.C., Ito,C., Csaszar,E., Roch,A., Yu,M., Sykes,E.A., Bader,G.D., and Zandstra,P.W. (2010). Dynamic interaction networks in a hierarchically organized tissue. Mol. Syst. Biol. *6*, 417.

Kitano,H. (2007). A robustness-based approach to systems-oriented drug design. Nat Rev Drug Discov. *6*, 202-210.

Kostek,M.C. and Delmonico,M.J. (2011). Age-Related Changes in Adult Muscle Morphology. Curr. Aging Sci.

Krishnamurthy,J., Torrice,C., Ramsey,M.R., Kovalev,G.I., Al-Regaiey,K., Su,L., and Sharpless,N.E. (2004). Ink4a/Arf expression is a biomarker of aging. J. Clin. Invest *114*, 1299-1307.

Kuhn,U., Gundel,M., Knoth,A., Kerwitz,Y., Rudel,S., and Wahle,E. (2009). Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J. Biol. Chem. *284*, 22803-22814.

Lage,K., Mollgard,K., Greenway,S., Wakimoto,H., Gorham,J.M., Workman,C.T., Bendsen,E., Hansen,N.T., Rigina,O., Roque,F.S., Wiese,C., Christoffels,V.M., Roberts,A.E., Smoot,L.B., Pu,W.T., Donahoe,P.K., Tommerup,N., Brunak,S., Seidman,C.E., Seidman,J.G., and Larsen,L.A. (2010). Dissecting spatio-temporal protein networks driving human heart development and related disorders. Mol. Syst. Biol. *6*, 381.

Lee,C.K., Klopp,R.G., Weindruch,R., and Prolla,T.A. (1999). Gene expression profile of aging and its retardation by caloric restriction. Science *285*, 1390-1393.

Lemay,J.F., D'Amours,A., Lemieux,C., Lackner,D.H., St-Sauveur,V.G., Bahler,J., and Bachand,F. (2010). The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. Mol. Cell *37*, 34-45.

Lemieux,C. and Bachand,F. (2009). Cotranscriptional recruitment of the nuclear poly(A)-binding protein Pab2 to nascent transcripts and association with translating mRNPs. Nucleic Acids Res. *37*, 3418-3430.

Lexell,J., Taylor,C.C., and Sjostrom,M. (1988). What is the cause of the ageing atrophy? Total number, size and proportion of different fiber types studied in whole vastus lateralis muscle from 15- to 83-year-old men. J. Neurol. Sci. *84*, 275-294.

Lindle,R.S., Metter,E.J., Lynch,N.A., Fleg,J.L., Fozard,J.L., Tobin,J., Roy,T.A., and Hurley,B.F. (1997). Age and gender comparisons of muscle strength in 654 women and men aged 20-93 yr. J. Appl. Physiol *83*, 1581-1587.

Linoli,G., Tomelleri,G., and Ghezzi,M. (1991). Oculopharyngeal muscular dystrophy. Description of a case with involvement of the central nervous system]. Pathologica *83*, 325-334.

Lipkowitz,S. and Weissman,A.M. (2011). RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. Nat Rev Cancer *11*, 629-643.

Liu,Z., Miers,W.R., Wei,L., and Barrett,E.J. (2000). The ubiquitin-proteasome proteolytic pathway in heart vs skeletal muscle: effects of acute diabetes. Biochem. Biophys. Res. Commun. *276*, 1255-1260.

Lu,T., Pan,Y., Kao,S.Y., Li,C., Kohane,I., Chan,J., and Yankner,B.A. (2004). Gene regulation and DNA damage in the ageing human brain. Nature *429*, 883-891.

Lu,Y., Huggins,P., and Bar-Joseph,Z. (2009). Cross species analysis of microarray expression data. Bioinformatics. *25*, 1476-1483.

Luscombe,N.M., Babu,M.M., Yu,H., Snyder,M., Teichmann,S.A., and Gerstein,M. (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. Nature *431*, 308-312.

Martin,G.M. (2009). Epigenetic gambling and epigenetic drift as an antagonistic pleiotropic mechanism of aging. Aging Cell *8*, 761-764.

Medici,M., Pizzarossa,C., Skuk,D., Yorio,D., Emmanuelli,G., and Mesa,R. (1997). Oculopharyngeal muscular dystrophy in Uruguay. Neuromuscul. Disord. *7 Suppl 1*, S50-S52.

Meusser,B., Hirsch,C., Jarosch,E., and Sommer,T. (2005). ERAD: the long road to destruction. Nat Cell Biol. *7*, 766-772.

Millefiorini,M. and Filippini,C. (1967). Oculopharyngeal muscular dystrophy. Riv. Neurol. *37*, 327-337.

Mizoi,Y., Yamamoto,T., Minami,N., Ohkuma,A., Nonaka,I., Nishino,I., Tamura,N., Amano,T., and Araki,N. (2011). Oculopharyngeal muscular dystrophy associated with dementia. Intern. Med. *50*, 2409-2412.

Moore,M.J. and Silver,P.A. (2008). Global analysis of mRNA splicing. RNA. *14*, 197-203.

Morimoto,R.I. (2008). Proteotoxic stress and inducible chaperone networks in neurodegenerative disease and aging. Genes Dev. *22*, 1427-1438.

Munch,C. and Bertolotti,A. (2010). Exposure of hydrophobic surfaces initiates aggregation of diverse ALS-causing superoxide dismutase-1 mutants. J. Mol. Biol. *399*, 512-525.

Nalepa,G., Rolfe,M., and Harper,J.W. (2006). Drug discovery in the ubiquitin-proteasome system. Nat Rev Drug Discov. *5*, 596-613.

Nicholson,J.K. and Wilson,I.D. (2003). Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. Nat Rev Drug Discov. *2*, 668-676.

Oberdoerffer,P. and Sinclair,D.A. (2007). The role of nuclear architecture in genomic instability and ageing. Nat Rev Mol. Cell Biol. *8*, 692-702.

Ochs,M.F. (2010). Knowledge-based data analysis comes of age. Brief. Bioinform. *11*, 30-39.

Oliva,A., Rosebrock,A., Ferrezuelo,F., Pyne,S., Chen,H., Skiena,S., Futcher,B., and Leatherwood,J. (2005). The cell cycle-regulated genes of Schizosaccharomyces pombe. PLoS. Biol. *3*, e225.

Olzmann,J.A., Li,L., Chudaev,M.V., Chen,J., Perez,F.A., Palmiter,R.D., and Chin,L.S. (2007). Parkin-mediated K63-linked polyubiquitination targets misfolded DJ-1 to aggresomes via binding to HDAC6. J. Cell Biol. *178*, 1025-1038.

Palla,G., Derenyi,I., Farkas,I., and Vicsek,T. (2005). Uncovering the overlapping community structure of complex networks in

nature and society. Nature *435*, 814-818.

Pan,Q., Shai,O., Lee,L.J., Frey,B.J., and Blencowe,B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. *40*, 1413-1415.

Pe'er,D., Regev,A., and Tanay,A. (2002). Minreg: inferring an active regulator set. Bioinformatics. *18 Suppl 1*, S258-S267.

Perie,S., Mamchaoui,K., Mouly,V., Blot,S., Bouazza,B., Thornell,L.E., St Guily,J.L., and Butler-Browne,G. (2006). Premature proliferative arrest of cricopharyngeal myoblasts in oculo-pharyngeal muscular dystrophy: Therapeutic perspectives of autologous myoblast transplantation. Neuromuscul Disord *16*, 770-781.

Pogribny,I.P., Ross,S.A., Tryndyak,V.P., Pogribna,M., Poirier,L.A., and Karpinets,T.V. (2006). Histone H3 lysine 9 and H4 lysine 20 trimethylation and the expression of Suv4-20h2 and Suv-39h1 histone methyltransferases in hepatocarcinogenesis induced by methyl deficiency in rats. Carcinogenesis *27*, 1180-1186.

Potapov,A.P., Goemann,B., and Wingender,E. (2008). The pairwise disconnectivity index as a new metric for the topological analysis of regulatory networks. BMC. Bioinformatics. *9*, 227.

Powers,E.T., Morimoto,R.I., Dillin,A., Kelly,J.W., and Balch,W.E. (2009). Biological and chemical approaches to diseases of proteostasis deficiency. Annu. Rev Biochem. *78*, 959-991.

Rajan,R.S., Illing,M.E., Bence,N.F., and Kopito,R.R. (2001). Specificity in intracellular protein aggregation and inclusion body formation. Proc. Natl. Acad. Sci. U. S. A *98*, 13060-13065.

Rakyan,V.K., Down,T.A., Maslau,S., Andrew,T., Yang,T.P., Beyan,H., Whittaker,P., McCann,O.T., Finer,S., Valdes,A.M., Leslie,R.D., Deloukas,P., and Spector,T.D. (2010). Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. *20*, 434-439.

Rando,T.A. and Chang,H.Y. (2012). Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. Cell *148*, 46-57.

Raz,V., Routledge,S., Venema,A., Buijze,H., van der Wal,E., Anvar,S.Y., Straasheijm,K.R., Klooster,R., Antoniou,M., and van der Maarel,S.M. (2011). Modeling Oculopharyngeal Muscular Dystrophy in Myotube Cultures Reveals Reduced Accumulation of Soluble Mutant PABPN1 Protein. Am. J. Pathol.

Reshef,D.N., Reshef,Y.A., Finucane,H.K., Grossman,S.R., McVean,G., Turnbaugh,P.J., Lander,E.S., Mitzenmacher,M., and Sabeti,P.C. (2011). Detecting novel associations in large data sets. Science *334*, 1518-1524.

Reverter,A., Ingham,A., and Dalrymple,B.P. (2008). Mining tissue specificity, gene connectivity and disease association to reveal a set of genes that modify the action of disease causing genes. BioData. Min *1*, 8.

Ron,D. and Walter,P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol. Cell Biol. *8*, 519-529.

Ronn,T., Poulsen,P., Hansson,O., Holmkvist,J., Almgren,P., Nilsson,P., Tuomi,T., Isomaa,B., Groop,L., Vaag,A., and Ling,C. (2008). Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. Diabetologia *51*, 1159-1168.

Roth,S.M., Ferrell,R.E., Peters,D.G., Metter,E.J., Hurley,B.F., and Rogers,M.A. (2002). Influence of age, sex, and strength training on human muscle gene expression determined by microarray. Physiol Genomics *10*, 181-190.

Saeed,R. and Deane,C.M. (2006). Protein protein interactions, evolutionary rate, abundance and age. BMC. Bioinformatics. *7*, 128.

Sahin,E. and Depinho,R.A. (2010). Linking functional decline of telomeres, mitochondria and stem cells during ageing. Nature *464*, 520-528.

Sandri,M. (2008). Signaling in muscle atrophy and hypertrophy. Physiology. (Bethesda.) *23*, 160-170.

Sarkar,A.K., Biswas,S.K., Ghosh,A.K., Mitra,P., Ghosh,S.K., and Mathew,J. (1995). Oculopharyngeal muscular dystrophy. Indian J. Pediatr. *62*, 496-498.

Schieppati,A., Henter,J.I., Daina,E., and Aperia,A. (2008). Why rare diseases are an important medical and social issue. Lancet *371*, 2039-2041.

Schilling,C.H., Edwards,J.S., and Palsson,B.O. (1999). Toward metabolic phenomics: analysis of genomic data using flux balances. Biotechnol. Prog. *15*, 288-295.

Schuldiner,M., Metz,J., Schmid,V., Denic,V., Rakwalska,M., Schmitt,H.D., Schwappach,B., and Weissman,J.S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cell *134*, 634-645.

Searls,D.B. (2005). Data integration: challenges for drug discovery. Nat Rev Drug Discov. *4*, 45-58.

Segal,E., Shapira,M., Regev,A., Pe'er,D., Botstein,D., Koller,D., and Friedman,N. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet. *34*, 166-176.

Segre,D., Deluna,A., Church,G.M., and Kishony,R. (2005). Modular epistasis in yeast metabolism. Nat Genet. *37*, 77-83.

Sexton,T., Schober,H., Fraser,P., and Gasser,S.M. (2007). Gene regulation through nuclear organization. Nat Struct. Mol. Biol. *14*, 1049-1055.

Soule,M. (1967). Phenetics of Natural Populations I. Phenetic Relationships of Insular Populations of the Side-Blotched Lizard. Evolution *21*, 584-591.

Spector,D.L. and Gasser,S.M. (2003). A molecular dissection of nuclear function. Conference on the dynamic nucleus: questions and implications. EMBO Rep. *4*, 18-23.

Spiro,Z., Kovacs,I.A., and Csermely,P. (2008). Drug-therapy networks and the prediction of novel drug targets. J. Biol. *7*, 20.

Sprinzak,E., Sattath,S., and Margalit,H. (2003). How reliable are experimental protein-protein interaction data? J. Mol. Biol. *327*, 919-923.

St Onge,R.P., Mani,R., Oh,J., Proctor,M., Fung,E., Davis,R.W., Nislow,C., Roth,F.P., and Giaever,G. (2007). Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. Nat Genet. *39*, 199-206.

Steele,E., Tucker,A., 't Hoen,P.A., and Schuemie,M.J. (2009). Literature-based priors for gene regulatory networks. Bioinformatics. *25*, 1768-1774.

Stolk,P., Willemen,M.J., and Leufkens,H.G. (2006). Rare essentials: drugs for rare diseases as essential medicines. Bull. World Health Organ *84*, 745-751.

Taillandier,D., Combaret,L., Pouch,M.N., Samuels,S.E., Bechet,D., and Attaix,D. (2004). The role of ubiquitin-proteasome-dependent proteolysis in the remodelling of skeletal muscle. Proc. Nutr. Soc. *63*, 357-361.

Tanay,A., Sharan,R., Kupiec,M., and Shamir,R. (2004). Revealing modularity and organization in the yeast molecular network by integrated analysis of highly heterogeneous genomewide data. Proc. Natl. Acad. Sci. U. S. A *101*, 2981-2986.

Tazi,J., Bakkour,N., and Stamm,S. (2009). Alternative splicing and disease. Biochim. Biophys. Acta *1792*, 14-26.

Teschendorff,A.E., Menon,U., Gentry-Maharaj,A., Ramus,S.J., Weisenberger,D.J., Shen,H., Campan,M., Noushmehr,H., Bell,C.G., Maxwell,A.P., Savage,D.A., Mueller-Holzner,E., Marth,C., Kocjan,G., Gayther,S.A., Jones,A., Beck,S., Wagner,W., Laird,P.W., Jacobs,I.J., and Widschwendter,M. (2010). Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. *20*, 440-446.

Tibshirani,R. (1996). Regression Shrinkage and Selection via the LASSO. Journal of the Royal Statistical Society *58*, 267-288.

Tohgi,H., Utsugisawa,K., Nagane,Y., Yoshimura,M., Genda,Y., and Ukitsu,M. (1999). Reduction with age in methylcytosine in the promoter region -224 approximately -101 of the amyloid precursor protein gene in autopsy human cortex. Brain Res. Mol. Brain Res. *70*, 288-292.

Tong,A.H., Lesage,G., Bader,G.D., Ding,H., Xu,H., Xin,X., Young,J., Berriz,G.F., Brost,R.L., Chang,M., Chen,Y., Cheng,X., Chua,G., Friesen,H., Goldberg,D.S., Haynes,J., Humphries,C., He,G., Hussein,S., Ke,L., Krogan,N., Li,Z., Levinson,J.N., Lu,H., Menard,P., Munyana,C., Parsons,A.B., Ryan,O., Tonikian,R., Roberts,T., Sdicu,A.M., Shapiro,J., Sheikh,B., Suter,B., Wong,S.L., Zhang,L.V., Zhu,H., Burd,C.G., Munro,S., Sander,C., Rine,J., Greenblatt,J., Peter,M., Bretscher,A., Bell,G., Roth,F.P., Brown,G.W., Andrews,B., Bussey,H., and Boone,C. (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Trinkle-Mulcahy,L. and Lamond,A.I. (2007). Toward a high-resolution view of nuclear dynamics. Science *318*, 1402-1407.

Trollet,C., Anvar,S.Y., Venema,A., Hargreaves,I.P., Foster,K., Vignaud,A., Ferry,A., Negroni,E., Hourde,C., Baraibar,M.A., 't Hoen,P.A., Davies,J.E., Rubinsztein,D.C., Heales,S.J., Mouly,V., van der Maarel,S.M., Butler-Browne,G., Raz,V., and Dickson,G. (2010). Molecular and phenotypic characterization of a mouse model of oculopharyngeal muscular dystrophy reveals severe muscular atrophy restricted to fast glycolytic fibres. Hum. Mol. Genet.

Tryndyak,V.P., Kovalchuk,O., and Pogribny,I.P. (2006). Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins. Cancer Biol. Ther. *5*, 65-70.

Tyedmers,J., Mogk,A., and Bukau,B. (2010). Cellular strategies for controlling protein aggregation. Nat Rev Mol. Cell Biol. *11*, 777-788.

Venkatesan,K., Rual,J.F., Vazquez,A., Stelzl,U., Lemmens,I., Hirozane-Kishikawa,T., Hao,T., Zenkner,M., Xin,X., Goh,K.I., Yildirim,M.A., Simonis,N., Heinzmann,K., Gebreab,F., Sahalie,J.M., Cevik,S., Simon,C., de Smet,A.S., Dann,E., Smolyar,A., Vinayagam,A., Yu,H., Szeto,D., Borick,H., Dricot,A., Klitgord,N., Murray,R.R., Lin,C., Lalowski,M., Timm,J., Rau,K., Boone,C., Braun,P., Cusick,M.E., Roth,F.P., Hill,D.E., Tavernier,J., Wanker,E.E., Barabasi,A.L., and Vidal,M. (2009). An empirical framework for binary interactome mapping. Nat Methods *6*, 83-90.

Vignaud,A., Ferry,A., Huguet,A., Baraibar,M., Trollet,C., Hyzewicz,J., Butler-Browne,G., Puymirat,J., Gourdon,G., and Furling,D. (2010). Progressive skeletal muscle weakness in transgenic mice expressing CTG expansions is associated with the activation of the ubiquitin-proteasome pathway. Neuromuscul. Disord. *20*, 319-325.

Wahle,E. (1991). A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. Cell *66*, 759-768.

Wahle,E. (1995). Poly(A) tail length control is caused by termination of processive synthesis. J. Biol. Chem. *270*, 2800-2808.

Wang,E.T., Sandberg,R., Luo,S., Khrebtukova,I., Zhang,L., Mayr,C., Kingsmore,S.F., Schroth,G.P., and Burge,C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang,G.S. and Cooper,T.A. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet. *8*, 749-761.

Wang,J., Farr,G.W., Zeiss,C.J., Rodriguez-Gil,D.J., Wilson,J.H., Furtak,K., Rutkowski,D.T., Kaufman,R.J., Ruse,C.I., Yates,J.R., III, Perrin,S., Feany,M.B., and Horwich,A.L. (2009). Progressive aggregation despite chaperone associations of a mutant SOD1-YFP in transgenic mice that develop ALS. Proc. Natl. Acad. Sci. U. S. A *106*, 1392-1397.

Wunderlich,Z. and Mirny,L.A. (2006). Using the topology of metabolic networks to predict viability of mutant strains. Biophys. J. *91*, 2304-2311.

Xu,G. and Jaffrey,S.R. (2011). The new landscape of protein ubiquitination. Nat Biotechnol. *29*, 1098-1100.

Yildirim,M.A., Goh,K.I., Cusick,M.E., Barabasi,A.L., and Vidal,M. (2007). Drug-target network. Nat Biotechnol. *25*, 1119-1126.

Yu,H., Kim,P.M., Sprecher,E., Trifonov,V., and Gerstein,M. (2007). The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. PLoS. Comput. Biol. *3*, e59.

Zahn,J.M., Sonu,R., Vogel,H., Crane,E., Mazan-Mamczarz,K., Rabkin,R., Davis,R.W., Becker,K.G., Owen,A.B., and Kim,S.K. (2006). Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS. Genet. *2*, e115.

Zhang,C., Frias,M.A., Mele,A., Ruggiu,M., Eom,T., Marney,C.B., Wang,H., Licatalosi,D.D., Fak,J.J., and Darnell,R.B. (2010). Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. Science *329*, 439-443.

Zhang,L.V., King,O.D., Wong,S.L., Goldberg,D.S., Tong,A.H., Lesage,G., Andrews,B., Bussey,H., Boone,C., and Roth,F.P. (2005). Motifs, themes and thematic maps of an integrated Saccharomyces cerevisiae interaction network. J. Biol. *4*, 6.

Zhou,X.J. and Gibson,G. (2004). Cross-species comparison of genome-wide expression patterns. Genome Biol. *5*, 232.