



# Comparative analyses of nuclear proteome: extending its function

**Kanika Narula, Asis Datta, Niranjana Chakraborty and Subhra Chakraborty\***

National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India

## Edited by:

Nicolas L. Taylor, The University of Western Australia, Australia

## Reviewed by:

Paula Casati, Centro de Estudios Fotosintéticos-Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Zhaohua Peng, Mississippi State University, USA

## \*Correspondence:

Subhra Chakraborty, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi-110067, India.

e-mail: subhrac@hotmail.com

Organelle proteomics is an emerging technology that is critical in determining the cellular signal transduction pathways. Nucleus, the regulatory hub of the eukaryotic cell is a dynamic system and a repository of various macromolecules that serve as modulators of such signaling that dictate cell fate decisions. Nuclear proteins (NPs) are predicted to comprise about 10–20% of the total cellular proteins, suggesting the involvement of the nucleus in a number of diverse functions. Indeed, NPs constitute a highly organized but complex network that plays diverse roles during development and physiological processes. In plants, relatively little is known about the nature of the molecular components and mechanisms involved in coordinating NP synthesis, their action and function. Proteomic study holds promise to understand the molecular basis of nuclear function using an unbiased comparative and differential approach. We identified a few hundred proteins that include classical and non-canonical nuclear components presumably associated with variety of cellular functions impinging on the complexity of nuclear proteome. Here, we review the nuclear proteome based on our own findings, available literature, and databases focusing on detailed comparative analysis of NPs and their functions in order to understand how plant nucleus works. The review also sheds light on the current status of plant nuclear proteome and discusses the future prospect.

**Keywords:** nucleus, comparative proteome, organelle proteome, plant, nuclear proteins

## INTRODUCTION

Cell nucleus has a perplex, heterogeneous, self-renewable, and dynamic social milieu which can sense signals, deformations, mechano-transduction, biochemical deliberations and many other processes ensuing outside its boundary. Nucleus is enclosed in a phospholipid rich membrane, which has very sensitive ion channels and pores that shuttle biomolecules in and out by conformational and morphological transformations. The nucleus, often referred as the “eukarya,” is functionally divided into the nuclear interior, a structurally differentiated and articulated organization surrounded by an envelope which is dynamic but sensitive to outside milieu (Dahl et al., 2008). In lieu, the modular disposition of indispensable, dynamic, and complex morphological features of nuclear locale administers its function. It paves the way for the role of nuclear architecture in critical regulatory processes. The nucleus is a fundamental component of the microenvironment of both plant and animal cells that has been substantially expanded during evolution and keeps the genetic material separate from other activities of cell (Roix and Misteli, 2002). It is the ultimate exhilaration of gene regulation and proteins directly controlling the gene expression (Wilson and Dawson, 2011). Furthermore, it has been reported that the nucleus plays an important morpho-regulatory role during organogenesis in animal, besides its pivotal role in chloroplast division in plants (Cavalier-Smith, 2006; Dahl et al., 2008). The plant nucleus has biomechanical and morphogenetic functions; it is a viscoelastic solid encompassing tenacity of protein complexes. The organization of nuclear proteins (NPs) into versatile assemblies provides precise control over the shape, size, and

composition of the nucleus, which opens a route toward the construction of sensors, programmable packaging and cargo delivery system within the sub-nuclear compartments as well as between the organelle. Plasticity in the nucleus allows cell differentiation, while rigidity in the nucleus determines its mechanical stiffness (Jiang et al., 2006; Pajerowski et al., 2007). Beyond its paramount importance in the generation of form, nucleus is frequently considered “growth-regulating” (Cavalier-Smith, 2006). The nucleus is evolutionary and inherently bestowed with information that can be both stored and relayed to cell interior via templating processes. It serves as the regulator in cell signaling for perceiving and transmitting extra- and inter-cellular signals in many cellular pathways. Communication between the cytoplasm and the nucleus is necessary and evident because of events such as apoptosis (Broers et al., 2002), mechanical stress (Dahl et al., 2008), environmental perturbation (Cheung and Reddy, 2012) and pathogen infection (Rivas, 2012), which lead to altered biosynthesis and modification of nuclear architecture and downstream cytoplasmic events. In addition, nucleoskeleton acts as a substrate for genome partitioning during mitosis. Further, it has been recognized as a central portal for providing motor centers during chromosome segregation in cell division (De Souza and Osmani, 2009). However, the available data is rather scarce and motor proteins between nucleoskeleton and chromosomes are still not known in higher plants. Throughout the plant kingdom the formation and regulation of the nuclear architecture has been shown to have the potential to influence many conduits of development, epigenetic differentiation, microfabricated patterning and cell senescence,

besides environmental stress response and pathobiology (Vergnes et al., 2004; Constantinescu et al., 2006; Dahl et al., 2008). Also, the nucleus serves a multi-functional role, as a regulator and modulator during cell division, and controller and integrator for fertilization and inheritance. Thus, nucleus plays a critical role as a modulator of cellular phenotype (Franklin et al., 2011). The nucleus must therefore be dynamic as cells divide, modulating its composition and architecture during its formation and after it has been disintegrated. The nuclear function is a multi-step, complex process, and the underlying mechanisms governing these steps are not fully understood.

All eukaryotic lineages are characterized by the loss, gain, expansion, and diversification of gene families (Fritz-Laylin et al., 2010). Understanding protein diversity and shared features can give unprecedented insight into the most fundamental aspects of nuclear structure and protein organization in as diverse kingdoms as plants and animals. Determination of organellar proteomes – the complement of proteins that reside, even if temporarily, in a specific organelle or sub-cellular region is of fundamental importance. Sub-cellular fractionation of tissue and cells in combination with MS/MS analysis has proven to be a powerful approach for the identification of proteins contained in specific organelles, such as the nucleus. Proteome research holds the promise of understanding the molecular basis of the nuclear function using an unbiased comparative and differential approach. Although the field of angiosperm eukaryogenesis has plethora of contradictory ideas, the nature of molecular changes can be reflected from the proteome. The nuptials of proteomics with cell biology have produced extensive inventories of the proteins that inhabit several sub-cellular organelles, including nucleus (Schirmer and Gerace, 2005; Yates et al., 2005; Yan et al., 2008). We and others have identified several hundred plant and animal NPs that include both predicted and non-canonical candidates, presumably associated with a variety of functions; viz., nucleoskeleton structure, development, DNA replication/repair, chromatin assembly/remodeling, signal transduction, mRNA processing, protein folding, transcription and splicing regulation, transport, metabolism, cell defense and rescue; all of which impinge on the complexity of NPs in plant (Pandey et al., 2006; Choudhary et al., 2009) and animal (Henrich et al., 2007). In recent years, reports have also been published focusing on changes in the nuclear proteome in varied cellular events (Bae et al., 2003; Lee et al., 2006; Salzano et al., 2006; Henrich et al., 2007; Buhr et al., 2008; Pandey et al., 2008; Repetto et al., 2008, 2012; Choudhary et al., 2009; Abdalla et al., 2010; Cooper et al., 2011; Varma and Mishra, 2011; Abdalla and Rafudeen, 2012). The identified proteins revealed the presence of complex regulatory networks that function in this organelle. NPs have been shown to account for approximately one-fourth of total proteins in yeast (Moriguchi et al., 2005) and one-fifth in animals (Bickmore and Sutherland, 2002), but the arithmetic estimate in plants is not yet complete. Currently, the focus is on nuclear proteomes in order to understand the nucleus-related processes in plants and animals. Although over the past few years there have been rapid advances in nuclear proteome research, the study on the complexity of NPs remained secondary, despite the fact they correspond to about 10–20% of the total cellular proteins and are comprised of several hundred different molecules with diverse functions. Moreover, a

vast array of post-translational modifications to these proteins add diversity to the structure and ligand-binding properties of nuclear components, leading to their differential activity. Therefore, characterization of the nuclear proteome in plant hold the promise of increasing our understanding about the regulation of genes and their function.

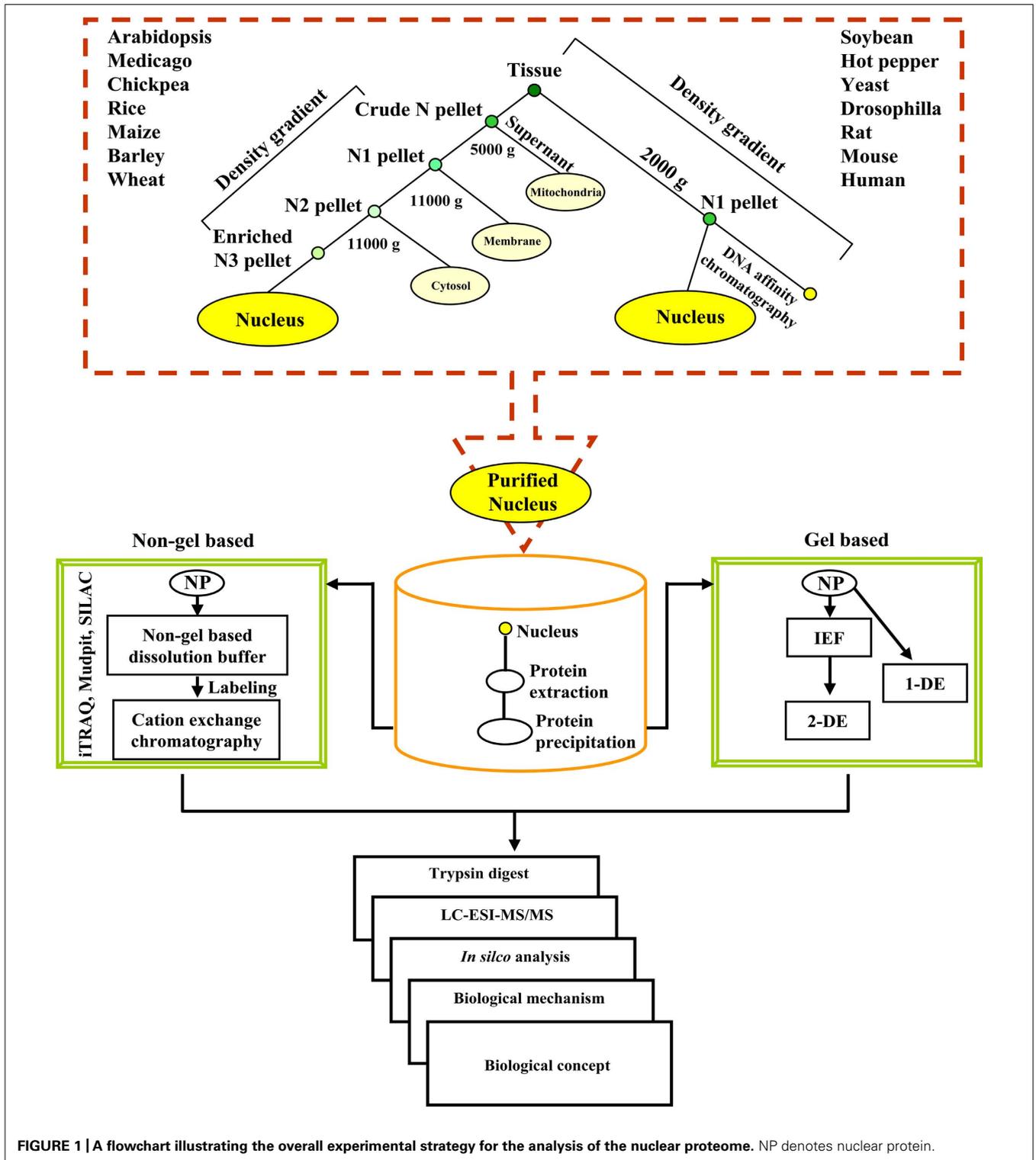
Here, we begin by giving updates on the nuclear proteomes and summarizing the essential and unique features of the nucleus. We also discuss recent findings concerning the regulation and biochemistry of it with specific emphasis on the fundamental role of NPs in development, DNA replication/repair, transcriptional regulation, environmental stress, and signaling by analyzing the nuclear proteomes. Furthermore, we report the cross-kingdom comparative analysis of nuclear proteomes toward organism specificity and plant exclusivity based on our own findings, the available literature and databases focusing on NPs in view of the current understanding and perspectives of the nuclear functions.

## ORIGIN OF THE NUCLEUS

A landmark event in the evolution of eukaryote was the acquisition of nucleus. Eukarogenesis have evolved albeit independently in plants and animals. Although both are true eukaryotes they have different ancestors. The evolution from prokaryotes to eukaryotes was the most radical change in cell organization. It is known that evolution of complex characters typically involves preadaptation, radical mutational innovation, and different selective forces acting in succession (for review, see Cavalier-Smith, 2006). Physical and mutational mechanisms of origin of the nucleus are seldom considered beyond the longstanding assumption that it involved wrapping pre-existing end membranes around chromatin (Cavalier-Smith, 1988). Evolution of the nucleus starts approximately 850 Million years ago (Cavalier-Smith, 2002), but it was 1833 when Robert Brown discovered the nucleus and said “vim and vigor is sexless devoid of this facet” in a paper to the Linnean Society. Origin of nucleus requires understanding of co-evolution of different nuclear components and their functional interlinking into the fundamentally novel eukaryotic life style. There are two competing theories of eukaryotic evolution. According to the first theory, a subset of bacteria slowly developed nucleus, while in the other, eukaryotes came first, some of them then lost nucleus and gave rise to bacteria. But, woesean revolution highlights that eukaryotes came from archaeal stock. Since, eukaryotes contain both archaeal and bacterial genes and the division of labor arose from the ancient symbiotic partnership between them that gave rise to eukaryotic nucleus. A third option for the nuclear origin revolves around the viruses, but the supporting data are provocative, circumstantial, and controversial (for review, see Pennisi, 2004).

## DESCRIPTION OF TOOLS TO STUDY NUCLEAR PROTEOME

An outline of the procedure and the illustration of the data that can be generated with the methodology are shown in **Figure 1**. Each proteomic study is described through a simplified flowchart showing its different steps from experimental material to protein identification. As illustrated in **Figure 1**, density gradient methods can be used to prepare a nuclear fraction with or without DNA



affinity chromatography. The most efficient means to separate NPs are either two-dimensional gel electrophoresis or cation-exchange chromatography followed by elution of protein fractions with salt gradient. Over the past few years 2-DE coupled MS/MS and LC-MS/MS have extensively been used to study nuclear proteomes

in varied organisms (for reviews, see Khan and Komatsu, 2004; Cullen and Mansuy, 2010; Erhardt et al., 2010). In brief, the NP fractions after separation are digested to allow identification of proteins by mass spectrometry. Proteins can be directly submitted to enzymatic digestion with appropriate proteases, such as trypsin

or to chemical treatment to get peptides of appropriate mass (usually between 750 and 4000 Da). Identification of proteins can then be done either by peptide sequencing using liquid chromatography coupled to MS (LC-MS/MS) or by peptide mass mapping using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF/TOF) followed by *in silico* analyses. Custom NP databases, for example, yeast-NPD<sup>1</sup>, human-NPD<sup>2</sup>, *Medicago*-NPD<sup>3</sup>, TAIR<sup>4</sup>, rice nuclear proteome database<sup>5</sup> help improve the identification and post-translational modification of the NPs. To undertake a comprehensive comparison of the plant nucleolar proteomes based on a combined approach of alignment, structure and phylogeny an *Arabidopsis* nucleolar protein database was curated (Brown et al., 2005). Similarly, comprehensive and well-annotated database of transcription factors may provide a useful resource to check annotations and to study gene regulatory pathways (Guo et al., 2005, 2008; Iida et al., 2005; Gao et al., 2006; Palaniswamy et al., 2006; Zhu et al., 2007; Rushton et al., 2008; Perez-Rodriguez et al., 2010; Romeuf et al., 2010; Wang et al., 2010).

Nuclear proteins are often under-represented in proteomic studies due to their low abundance. The information offered from total nuclear proteome by high-throughput techniques does not illustrate the functional purpose of NPs and compartment structures. Computational modeling, on the other hand, may elucidate functional roles otherwise not captured by any individual existing experimental technology. The predictions used to identify the common organelle-specific sequence features are successful for over-represented proteins but is limited for the low abundant proteins. Thus, this analysis allows the identification of additional proteins sharing the same motif and to estimate the enrichment of the protein motifs in nuclear proteome data set. The estimation of the enrichment of those motifs in the nuclear proteome data set was done by comparing their frequencies in the nuclear data sets and in the target protein databases. Some of the motifs that can be identified in the nucleus are domains of well-known proteins, including histones and helicases. There are many proteins that are known to be imported into the nucleus, but which have no known intra-nuclear compartment association. These proteins may share similar cellular locations or functions, but further experiments are needed for clarification (Gorski and Misteli, 2005).

## WHAT HAVE WE LEARNT?

Proteomics has turned out to be an imperative benefactor for studying the acquaintance of plant nuclear structure and function. The field of proteomics is evolving from cataloguing the proteins under static conditions to comparative analyses (Narula et al., 2012). Defining proteins that change in abundance, form, location or other activities may indicate the presence and functional significance of a protein. Whereas comparative nuclear proteome research is quite advanced in animals (Liao et al., 2009) and yeast (Gauci et al., 2009), there is less information in plants. The investigation on plant nuclear proteomes in recent years has

raised the following important questions: What are the essential plant NPs? Do NPs show clade specificity in vascular plants? What are those organ-specific NPs, if any? Does the nuclear developmental proteomics of one of the clades yield any astonishing or prolific results? How do NPs remodel during environmental- and/or patho-stress those provide new perspectives? Are some of the NPs unexpected? And, last but not the least, what sort of post-translational modifications have so far been characterized in the nucleus? Here, we analyze and compare the experimental results thus far available on nuclear proteomes to elucidate the dynamics of plant NPs.

## DECIPHERING THE ORGANISM-SPECIFIC NUCLEAR PROTEOME DYNAMICS: SOCIAL CLASS VS. DIVERSITY

Proteins evolve at rates differing over many orders of magnitude. As new proteins evolve by gene duplication, evolutionary rates must change dramatically over time. They change systematically among different branches of the evolutionary tree and also episodically (Cavalier-Smith, 2010). In the history of life there are three mega evolutions giving rise to prokaryotes, plants, and animals. Consequently, decoding organism specific nucleoid/nuclear proteomes are of utmost importance to understand the diversity among the protein complement in the three life lines of fruition. The possibility of intra-kingdom and/or cross-kingdom comparison of proteins and cellular regulation with the use of advanced proteomic techniques are of great value. We compared the experimentally determined nuclear proteomes of plants *viz.*, *Arabidopsis thaliana*, *Cicer arietinum*, *Medicago sativa*, *Glycine max*, *Capsicum frutescens*, *Xerophyta viscosa*, and *Oryza sativa* with that of yeast, fruit-fly and animal (human, rat, and mouse; **Table 1**). The *modus operandi* in investigating the nuclear proteomes of available species were the extensive literature search, availability of relevant databases (human, mouse, rat and plant, TIGR, UniProt and Swissprot) and *in silico* analysis. After comparing individual NPs, we investigated the divergence of these proteins among animals and plants to understand the integration and coordination of nuclear functions. Further, we calculated the percentage of proteins that was found to be unique to each proteome by calculating the number of proteins estimated from matches to SWISS-PROT as described in Skovgaard et al. (2001), Semple et al. (2003), and Bhushan et al. (2006). The NPs identified in these studies were classified into different functional categories. This classification is only tentative, since the biological role of many of the proteins identified has not been established experimentally. Furthermore, we applied a cross-species comparison on the available datasets. When analyzing proteomes within the specified group of plants, a logical strategy was used to maximize efficiency and the overall comparative results. Thus, it was imperative to first evaluate the available nuclear proteome maps, followed by an analysis of stimulus-specific proteomes of the above mentioned organisms. We then moved on to assess the stress-responsive plant nuclear proteomes in order to understand the overlap and specificity amongst different environmental- and patho-stress. These comparative studies were customized for specific protein families. It is to be noted that protein consensus can be obtained across any combination of proteomes based on the type of extraction procedure.

<sup>1</sup>www.pin.mskcc.org

<sup>2</sup>www.npd.hgu.mrc.ac

<sup>3</sup>www.masc.proteomics-org

<sup>4</sup>www.arabidopsis.org

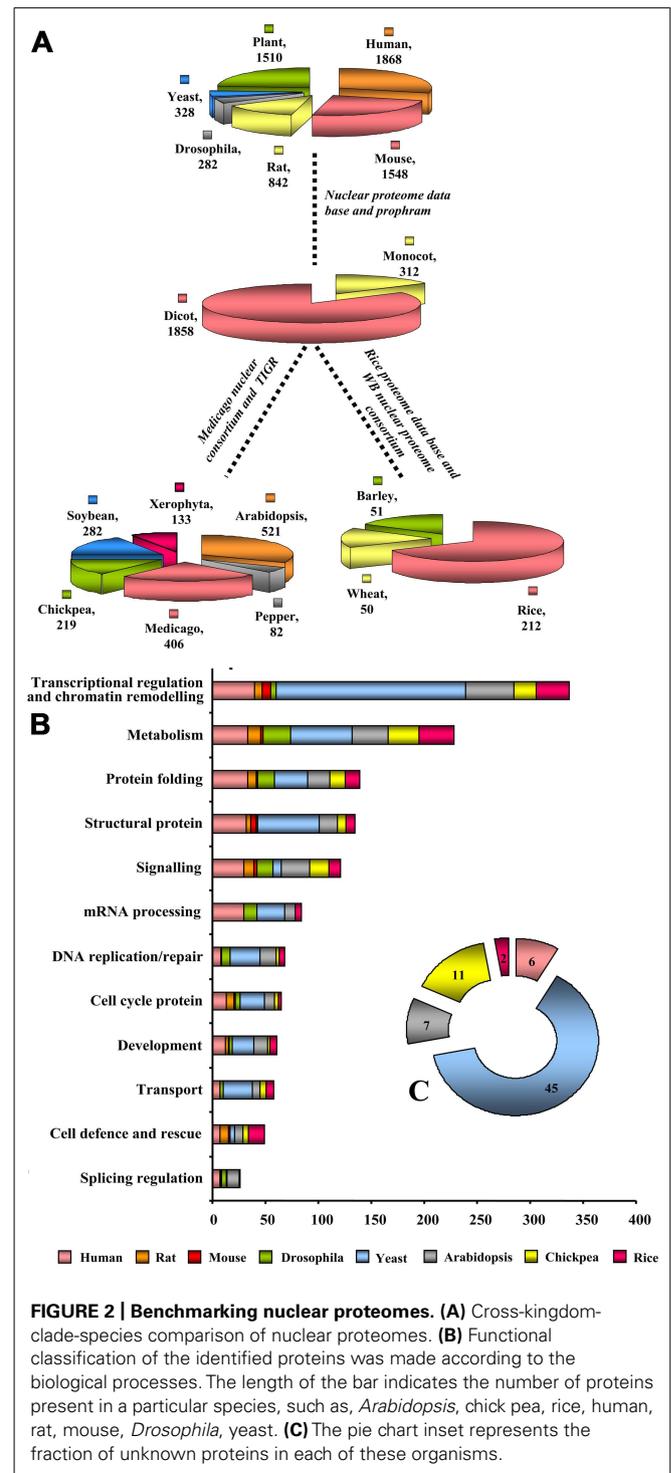
<sup>5</sup>http://gene64.dna.affrc.go.jp/RPD

Table 1 | A comprehensive list of plant nuclear proteomes.

S. No.	Organism	No. of proteins identified	Biological material used for fractionation	Technique used	Experimental conditions	Reference
1	<i>Chlamydomonas reinhardtii</i>	672	Cell suspension culture of algae	nanoLC-MS/MS	Cell	Winck et al. (2011)
2	<i>Arabidopsis thaliana</i>	154	Leaves	2DE/LC-MS/MS	cold	Bae et al. (2003)
3	<i>Medicago truncatula</i>	143	Seed	nano-liquid chromatography-nano-LC-MS/MS	Development	Repetto et al. (2008)
4	<i>Cicer arietinum</i>	150	Leaves	2DE/LC-MS/MS	Tissue	Pandey et al. (2006)
5	<i>Cicer arietinum</i>	150	Leaves	2DE/LC-MS/MS	Dehydration	Pandey et al., 2008
6	<i>Xerophyta viscosa</i>	18	Leaves	2-DE/ MALDI-TOF-TOF	Dehydration	Abdalla et al. (2010)
7	<i>Xerophyta viscosa</i>	121	Leaves	2-DE/ ESI- MS/MS	Dehydration	Abdalla and Rafudeen (2012)
8	<i>Glycine max</i>	292	Leaves	nano-liquid chromatography-nano-LC-MS/MS	Rust fungus	Cooper et al. (2011)
9	<i>Capsicum annuum</i>	56	Leaves	2-DE/MALDI-TOF-MS	Tobacco mosaic virus	Lee et al. (2006)
10	<i>Linum usitatissimum</i>	292	Immature seed coat at 16daf (Torpedo stage)	1-DE, WB, Dot blot, gel shift	Development	Renouard et al. (2012)
11	<i>Oryza sativa</i>	468	Endosperm	LC-MS/MS	Tissue	Li et al. (2008)
12	<i>Oryza sativa</i>	109	Leaves	2DE/LC-MS/MS	Dehydration	Choudhary et al. (2009)
13	<i>Oryza sativa</i>	307	Leaves	Percoll gradient centrifugation-nano LC-MS/MS	Glucose responsive	Aki and Yanagisawa (2009)
14	<i>Zea mays</i>	Not identified	Endosperm	1-DE, WB	Development	Ferreira et al. (2006)

Nuclear protein composition was found to differ between two major Kingdoms viz., plant and animal. Results were predominantly obtained with human (lymphoma and myeloid lines; Salzano et al., 2006; Henrich et al., 2007), mouse (Buhr et al., 2008), rat (McClatchy et al., 2011), *Drosophila* (Varma and Mishra, 2011), model plants *Arabidopsis* (Bae et al., 2003), *Medicago* (Repetto et al., 2008, 2012), and crop plants rice (Choudhary et al., 2009), hot pepper (Lee et al., 2006), soybean (Cooper et al., 2011), *Xerophyta* (Abdalla et al., 2010; Abdalla and Rafudeen, 2012), and chickpea (Pandey et al., 2006, 2008). The stunning findings from these comparisons suggest that until now only 1868 NPs are identified in humans, while 1548 in mouse, 842 in rats, 282 in *Drosophila*, 328 in yeast, and 1510 in plants contributing to the large repertoires of the nuclear proteome database and prophram (Figure 2A). A more accurate vision of animal nuclear proteome illustrate that approximately 2000 NPs from lymphoid and myeloid tissues of humans symbolizing around one-third of their estimated nuclear proteome. The two branches of angiosperm, monocot, and eudicot, firmly set up compelling evidence from the available nuclear proteomes that some NPs are unique, while some are shared. Extended nuclear proteome research is required for monocot family with only 312 NPs reported thus far than dicots that reported 1856 NPs (Figure 2A). Contextual information on nuclear proteomes of eudicots revealed that until now 521 NPs have been identified in *Arabidopsis* representing about one-third of its estimated nuclear proteome (Bae et al., 2003), while 406, 282, 219, 82, and 133 NPs were identified in *Medicago*, soybean, chickpea, hot pepper, and *Xerophyta*, respectively. To explore early messages arising from comparison of the content of monocot and dicot proteomes address key consequences of research for dicot comparative proteomics. The recurring observation that monocot proteome research centers on rice proves factual for nuclear proteome. Until recently, 212 NPs have been identified in rice, whereas only 50 and 51 NPs are known in wheat and barley, respectively (Figure 2A). Our comparative analyses of different species in relation to their function showed that high percentage of proteins to be unique to each proteome: 89% in animal (human, rat, and mouse), 81% in human (lymphoma and myeloid lines), 71% in mouse, 68% in rat, 84% in yeast, and 74% in *Drosophila*; whereas plant proteomes show 85% in *Arabidopsis*, 78% in soybean, 81% in chickpea, 71% in *Medicago*, 84% in rice, and 54% in hot pepper with only actin and 26S proteasome being the social class of proteins present ubiquitously in all. The available nuclear proteomes of nine plants compared in Figure 2A varied in molecular weight from 9.1 to 150 kDa and had a spread of *pI* values from 3.6 to 10; while yeast shows 15 to 110 kDa, 3.1 to 12.0 *pI*; *Drosophila* shows 12 to 140 kDa, 3.1 to 11.4 *pI* and animals show 9.4 to 150 kDa, 3.0 to 12.0 *pI*. Most of the NPs were basic in nature concordant with the acidic environment of this organelle.

Functional categorization of nuclear proteomes reported till date revealed an imperative corollary, which shows overall proteins belonging to transcriptional regulation and chromatin remodeling contribute radically to nuclear proteomes of yeast (54%), plants (29%), animals (16.5%), and *Drosophila* (1.2%). 54% of proteins in this category in yeast represent more than half of the total NP implying thereby that yeast is a dynamically dividing



**FIGURE 2 | Benchmarking nuclear proteomes. (A)** Cross-kingdom-clade-species comparison of nuclear proteomes. **(B)** Functional classification of the identified proteins was made according to the biological processes. The length of the bar indicates the number of proteins present in a particular species, such as, *Arabidopsis*, chick pea, rice, human, rat, mouse, *Drosophila*, yeast. **(C)** The pie chart inset represents the fraction of unknown proteins in each of these organisms.

organism having plurality of transcription regulators. Of these, 42% represents nuclear structural proteins involved in cell division. Furthermore, plants are deskbound, therefore to establish, maintain and alter global and local level of nucleic acid they require rapid turnover of DNA and RNA metabolizing proteins, DNA replication/repair proteins, and splicing regulation proteins toward acclimatization in the environment. Indeed, identified

plant NPs from the available reference proteomes showed 42% NPs belong to metabolism category whereas 46% confers splicing regulation. NPs behave like a network scaffold and acts as an entry point to ensure smoother regulation of different cellular processes that require rapid protein turn over. Comparison of plant nuclear proteomes with other organisms revealed that protein folding and turnover category contributes to 35%, which is in close correlation to 30% NPs from animals. Plant protein networks revealed the predominance of the development specific proteins (36%) and cell cycle proteins (24%). It is to be noted that the NP extraction protocol used in the animal and plant are different. Animal nuclear proteome research spotlight sub-nuclear compartmentalization, whereas, plant nuclear proteomes except *A. thaliana* (Calikowski et al., 2003; Pendle et al., 2005) and *O. sativa* (Tan et al., 2007, 2010) tranquil total nucleus. Therefore, the chromatin assembly/remodeling proteins are identified much less in all organisms as they can be isolated best in high salt buffer concentration which is not usually used for total NP extraction. A dramatic rearrangement of the nuclear structure takes place during mitosis and meiosis, which dynamically changes sub-nuclear proteomes (Beven et al., 1995; Holmes-Davis and Comai, 1998). Major role of interphase chromatin is in transcription, while mitotic chromatin contributes to cell division and meiotic chromatin is engaged in pairing, cross-over, and chromosome segregation. Compositional divergence in protein complement of stationary and dividing nucleus is thus a call to study proteomes at different phases of nuclear division.

Between plant and animal, the gene families and members are not related but functionally appear to be similar. Mosaic comparison of nuclear proteomes revealed that chimeric evolution was the main cause of proteome diversity in animals. For example, in human nucleophosmin protein has diverse protein members, whereas in mouse it is conserved. Likewise, when rat and mouse were compared DEAD, septin, lamin B2 box are some of the diverse class of proteins found in rodents. Also, when human, rodents and *Drosophila* were evaluated lactate dehydrogenase, nicotinamide synthase1, tubulin beta, and 26S proteasome represented the social class. Furthermore, when, human, rodent (mouse, rat), insect (*Drosophila*) and fungus (yeast) were compared actin and nicotinamide synthase 1 correspond to the social class. Whereas, comparison between monocots and dicots, showed that *Arabidopsis* is better explored than rice and therefore comparison of their proteomes may not yield the postulated results as defined by genome analysis. But the proteins revealed an evolutionary divergence in plant as well as dicot vs. monocot specificity, with few conserved proteins (**Figure 2B**). When the *Medicago* nuclear proteome was compared with that of *Arabidopsis*, results revealed an evolutionary divergence as well as tissue specificity, with few conserved proteins (**Figure 2B**). Comparison of the functional classes of NPs amongst dicot species like *Arabidopsis*, *Medicago*, soybean, hot pepper, *Xerophyta*, and chickpea confirm the dynamic and heterogeneous nature of nucleus as exemplified by the presence of only actin in all dicots. Another protein namely 26S proteasome may be considered as social class except its absence in *Xerophyta*. The presence of chaperone 60 in *Cicer* and HSP71in *Medicago* illustrate that nature invented vastly different solutions to a common problem viz., protein folding. When the studies

on the legumes like *Cicer* and *Medicago* were compared to *Arabidopsis* belonging to the Brassicaceae family (**Figure 2**), it can be readily observed that the splicing regulation in the nucleus for activating splicing enzymes is diverse between the two families as well as between the members of the same family, leguminosae. The protein network of rice revealed the predominance of the chromatin assembly/remodeling proteins, for example, histone deacetylase, histone 2A, histone 2B, histone 3, histone 4, while the *Arabidopsis* protein network was found to be rich in splicing regulation proteins and structural protein as transcriptional regulators.

It may be assumed that the divergence in the resulting proteomes of the vascular plants is due to the presence of the different nuclear architecture based on the protein and nucleic acid compositions, suggesting the occurrence of clade-specific NPs that would bind to their cognitive biomolecules to bring out specific functions both spatially and temporally. Most intriguing are the remaining 10–18% of plant NPs that do not have any similarity to the known proteins in other organisms. The challenge is to elucidate their biological role within the cell nucleus.

### EXPLORING THE SINK AND LINK IN NUCLEUS

Ubiquitously present, except in RBC, the nucleus is composed of different molecules with diverse functions to meet the specialized requirements of different organs and tissues. Nuclear functional compartmentalization is a paradigm of molecular machines necessary for biogenesis and functionality (Strouboulis and Wolffe, 1996). It is a dynamic milieu having a reservoir for bioactive molecules, such as carbohydrates, nucleic acid, and proteins which is necessary for assembly and also for communication with the other parts of the cell. For decades, cell nucleus has been a black box in biology. The determination of comprehensive chemical differences between plant and animal nucleus is still difficult to understand, but the switching of cellular programs by NPs mediated chemical networking is tightly linked to the regulation of gene expression in both the kingdoms. However, the distributions of transcription sites in chromosome territories are conserved in plants and animals. It is the heterochromatic centers which makes the difference in nuclear processes in both these kingdoms (van Driel and Fransz, 2004). Being a store house of nucleic acid and proteinaceous domain, nucleus contains distinct structural and functional compartments (Misteli, 2005). Proteinaceous domain include nucleolus containing rRNA binding protein and splicing proteins, the cajal body having snRNA forming and binding proteins; whereas the nucleic acid domain encompass euchromatin and heterochromatin. Euchromatin is the reservoir of histones and histone binding proteins, while heterochromatin consists of heterochromatin binding unknown proteins. In these two domains nucleic acid occurs three dimensionally (Dundr and Misteli, 2001). Nuclear bodies are functionally and/or morphologically discrete accommodating usually distinct resident proteins. Paradigm includes the nucleoli (site of rRNA transcription), nuclear speckles (site for splicing) and splicing factor compartment (store-house for cajal body and PML body; Takizawa and Meshorer, 2008). Dramatic developments in high-resolution live-cell imaging have revealed the cell nucleus as a highly heterogeneous and complex organelle, and the global genome and proteome architecture

changes during processes such as differentiation and development (Misteli, 2001; Spector, 2003; Misteli, 2005). It is, therefore, relevant that different family members show highly regulated and specific patterns of the expression of nuclear components in an evolutionary context. Similarities in nuclear design may be apparent as it is likely that ancient functional protein domains and nucleic acid backbones have been used in a variety of arrangements and combinations to affect the function of convergent biological structures. Nucleus serves as the self organizing mediator. Most proteins are in constant motion, and their residence time within a compartment is very low, being at most 1 min (Gonzalez-Melendi et al., 2000). This mobility ensures that proteins find their targets by energy-independent passive diffusion (Pederson, 2000). In addition to protein heterogeneity and the presence of various regulators, mediators, transducers as well as linkers, RNA and chromatin compositions can vary between cell types and even within a given cell in different time (Hetzer et al., 2005), suggesting that the nucleus serves as a sink of variability in terms of macromolecules or microelements. Regulated trafficking of proteins, RNAs, RNA-protein complexes, and other molecules in and out of the nucleus is important in diverse processes. The nucleus serves as the end line culminator in cell signaling to perceive and transmit extra- and intercellular signals in many cellular pathways. NPs not only constitute more than just a structural scaffold, but also play various roles in development, cell cycle, defense against environmental stresses and in the tight regulation of gene expression.

### THE NUCLEAR PROTEIN SINK: A DYNAMIC FRAMEWORK FOR MULTIPLE FUNCTIONS

Eukaryotic NPs are complex with plurifunctional role, evolutionary tinkering, and subtle modifications evoked repeatedly and independently among different taxa. A macromolecular machine in the form of nuclear pore allows a protein or protein complex up to approx. 500 kDa to traverse the nucleus. NPs roam through the nucleus in search of a high-affinity binding site where they can exert their functions (Dundr and Misteli, 2001; Mans et al., 2004). The specific domain and architecture of NP contain information of biological importance and evolutionary value.

Altogether, NPs include those which are highly mobile viz., transcription factors, pre-mRNA splicing factors, rRNA processing enzymes and 3 $\alpha$ -processing factors, DNA repair enzymes, chromatin-binding proteins and apoptotic caspases; while immobilized NPs encompass DNA replication factors, intermediate filament proteins, and histones H1 (Dundr and Misteli, 2001). Plant and animal show least homology as far as nuclear intermediate filament proteins are concerned. However, proteins belonging to DNA replication/repair are found to be orthologous (Moriguchi et al., 2005). In plants, elongation factor thermo unstable (EF-Tu), Zinc finger protein, glycine rich RNA binding, histone 2B, histone 3, glycine dehydrogenase, peptidyl prolyl isomerase, 26S proteasome, 60 kDa chaperone, glyceraldehyde 3-phosphate dehydrogenase, malate dehydrogenase, peroxiredoxin, transaldolase, calcium protein kinase, PHO1 like protein,  $\alpha$  expansin, actin, 14-3-3, and 40S ribosomal protein SA are consistently represented in thus far studied nuclear proteomes that play diverse

and crucial roles in nuclear function. Most predominant class of NPs reported are the transcription regulators in which TF2A, RNA polymerase have been optimized during eukaryotic evolution for acting in post-transcriptional gene regulation. The linear representation of promoter elements provides competency for physiological responsiveness within the contexts of development, cell cycle, and phenotype-dependent regulation as transcription factors can bind to these cis-acting elements dictating where and when a gene to be active. Chromatin binding proteins and nucleosome organization protein viz., PolII, MADS box, RCC2 protein, and HEAT box reduce distances between independent regulatory elements providing a basis for integrating components of transcriptional control. It is known that the nuclear matrix proteins support gene expression by imposing physical constraints on chromatin related to the three-dimensional genomic organization. In addition, the nuclear matrix proteins facilitate gene localization besides the concentration and targeting of transcription factors. Histone deacetylase 6 and DNA methyltransferase physically interact; together they mediate histone acetylation and modulate DNA methylation status, silencing the transposable element (Casati et al., 2008; Casati, 2012). Transcriptional reprogramming by WRKY, ERF, TGA, Whirly, and MYB factors is thought to cause alteration in transcript level, which in turn regulates various physiological processes like growth, development, and pathogen perturbation (Mayrose et al., 2006; Yu et al., 2011; Feng et al., 2012). Among others, it is ascertain that G5bf protein and TF rough sheath 2 are embodied persistently in dicot nuclear proteome in customary environment. Perhaps, the protein most expected to be similar to their metazoan counterpart in the plant nucleus is DNA ligase, which have been shown to regulate transcription (Truncaite et al., 2006). RF2B, SPT2-chromatin binding domain, RING zinc finger protein, and gypsy-like retroposon are exclusively present in monocot nuclear proteomes (Li et al., 2008; Aki and Yanagisawa, 2009; Choudhary et al., 2009). Aforesaid, transcription regulators of two clades have solitary similarity that they are regulated by circadian rhythm and have multivariate decision to find motif combination (Cavalier-Smith, 2010). Proteome data indicate that BABY BOOM, AP2/EBEBP2, and syringolide induced proteins are leguminosae allied transcriptional regulators having role in development, cell/organ identity and fate; while ribosomal recycling factor, CHP rich zinc finger protein, nucleolin, RuvB, BRI KD interacting protein, WPP domain protein, pescadillo protein, and MYB transcription factor are solanaceae associated transcriptional regulators (Boutilier et al., 2002; Abe et al., 2003). Each of these proteins reported in leguminosae and solanaceae have been shown to be involved in diverse cellular functions, viz. development, embryogenesis, and signaling pathways. This further highlights the technical challenges when attempting to isolate high purity nucleus and resolution of proteins using proteomic technology. Proteins involved in the metabolism are customary in the case of any nuclear proteome. Indeed, methylenetetrahydrofolate reductase, homocysteine methyl transferase, methyltransferase, and ornithine aminotransferase are proteins belonging to this category, which play pivotal role in RNA and DNA metabolism. Astounding result obtained from the comparative analysis of nuclear proteomes in plants with that of animals, suggests the

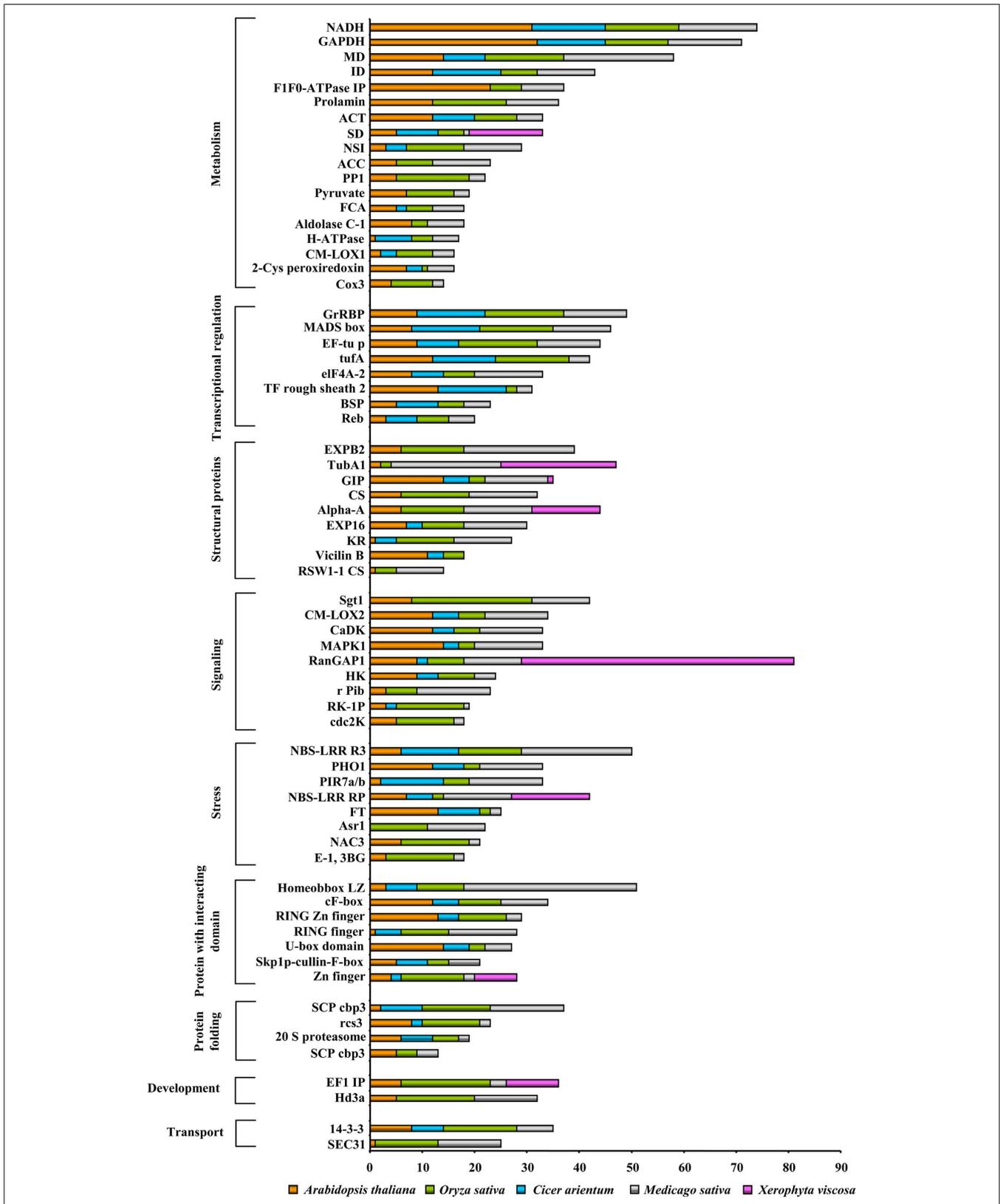
presence of many metabolism related animal orthologs in dicot (*Arabidopsis*) proteome whose role in plants have not yet been defined viz., biliverdin reductase A1, LROS1 acyl transferase, and KES1 oxysterol binding protein. Nuclear structural proteins are ubiquitous in both the kingdom but shows more divergence in plants (Meagher et al., 1999). Actin and myosin are the ancient component of this category that form a platform for all three DNA-dependent RNA polymerases, mediate RNA export from the nucleus, and are required for the long-range movement of specific loci within the nucleus (Caudron-Herger and Rippe, 2012). Recent evidence suggests that proteins such as actin, myosin, tubulin, NuMA, Annexin A1, Annexin A2, viscinalin, spectrins, and titin are recognized as having fundamental roles in nuclear structure and genome function in living eukaryotes (Wilson and Dawson, 2011). Coiled-coil protein, an orthologue of lamin, the building block of the NPC complex in plants (Mans et al., 2004) has many candidates, namely disease resistance proteins, NUF1, NUP82, NUP88. Nucleoporins, which anchor intermediate filament proteins during scaffold formation play a crucial role in chromosome scaffolding and mRNA export. The nucleoskeleton and nucleopore complex protein present ubiquitously in plants are expansins and NUPs. They maintain the mechanostatic and load bearing properties of the nucleus (Dahl et al., 2008). In other words, nucleostuctural dynamics in plant cell is a team effort of multiple proteins orchestrating this very fast-paced game. Nucleus is an evolutionary chimera of cell cycle related proteins. Progress made in the area of plant cell cycle regulation has resulted in recognition of NP candidates, including chromobox proteins, RCC2 proteins, and BUB3 having role in cell division. Additionally, CDC 5, one of the cell cycle proteins has a varied role in mitosis, ciliary motility and trafficking. Another component of cell cycle regulation namely, Ran cycle is represented by nuclear Ran GTPase, mago nashi protein, Ras GTPase, and Ran. Ribosome subunit export system of nucleus involved in cell cycle regulation focus on the three most appealing candidates: Nops, Nugs, GTPases, besides recently added AAA-ATPase and exportin in animals. However, analyses of plant nuclear proteomes do not show the presence of these proteins. Perhaps another protein most expected to be similar to their metazoan counterparts in the plant cell nucleus is karyopherins, which has a role in nuclear trafficking. Nucleus includes numerous enzymes viz., rRNA processing enzyme, polymerase, ligase, gyrase, and number of helicase that alter DNA conformation, replication, degradation, and chromatin modifiers. Histone variants, one of the important classes of NPs in eukaryotes are important component which play a key role in genome maintenance and stability. In the nucleus, RNA is an architectural factor for shaping the genome and its nuclear environment, besides being an effector molecule in maintaining the chromatin structure (Ma et al., 2011). We find many mRNA processing proteins that include nucleolar RNA-associated protein (NRAP), LSM2, paraspeckle protein 1, and Non-POU domain containing protein in this category. It is well known that protein folding supports diverse but specific signal transducers and lies at the interface of several developmental pathways (Caudron-Herger and Rippe, 2012). Likewise, different chaperones, HSP71, proteasome subunit alpha

types, DnaJ, protein disulfide isomerase, HSP20, glutathione-S-transferase, and HSP70 reported in plant nuclear proteomes might maintain protein homeostasis by providing stability to other nuclear resident proteins. Involvement of some of these chaperones with the class of developmental NPs viz., DEAD box, DUX3, von wilberand factor, HOME0 BOX, and U box have already been reported (Barthelery et al., 2008; Su and Li, 2008). A chronic theme proverbial to the class of nucleoskeleton linker proteins of plant cells is that these mechano-transducing transmembrane molecules communicate and interact preferentially with the intermediate filament on the nuclear side of the nuclear membrane. Our analyses suggest, several attributes of NP contribute to cross-talk in gene regulation and cellular phenotype.

### NUCLEAR INVENTORIES FOR *IN SILICO* PROTEIN PROFILING OF COMPARATIVE STRESS PROTEOME

Nucleus senses and physiologically responds to environmental stress via signaling pathways. Signaling events are clearly not linear and induce many different reactions, including stress-related processes that crosstalk with hormone signaling pathways. Most signaling pathways culminate in the nucleus leading to regulation of expression of specific genes whose products are necessary for eliciting a signal specific response like nuclear localization of pathogen effectors, R proteins, and other host defense proteins that modulate stress response. Here, we have customized the comparative analyses for specific protein families. For example, when the environmental stress-responsive proteomes were compared, the parallel analysis of the proteomes of different clades of vascular plants were performed, viz., chick pea vs. *Xerophyta* vs. rice for dehydration, *Arabidopsis* for cold response, and *Medicago* for seed filling that mimic the dehydration response. Similarly, in case of patho-stress, soybean, and hot-pepper proteomes were compared.

We analyzed the nuclear proteomes of *A. thaliana* in response to cold-stress (Bae et al., 2003) and dehydration-responsive nuclear proteomes of *Cicer arietinum* and *O. sativa* (Pandey et al., 2008; Choudhary et al., 2009). Interestingly, a great level of divergence in the protein classes amongst these organisms was observed (Figure 3). To our surprise, except development category all of the NPs were found to have members common in all organisms under both kind of abiotic stresses studied. Families of development related proteins, viz., embryonic flower 1- like protein, copia-like, and Hd3a protein have been found in dehydration responsive proteome of rice, while chickpea DRPs exclude most of the nuclear structural proteins such as cellulose synthase like, alpha amylase, and beta-expansin otherwise abundantly present in dehydration-responsive proteome of rice. It is intriguing to note that cold-responsive NPs under all functional categories of *Arabidopsis* were present in dehydration-responsive proteomes of rice and chickpea. Another important finding was the presence of cyc3 protein in high abundance during cold-stress in *Arabidopsis*. Whereas zinc finger, ring finger, and RNA glycine rich proteins were predominantly found during dehydration response but were absent in response to cold-stress. Various kinases known to mediate the stress-induced synthesis of NPs, such as PHO1, galectin, thioredoxin peroxidase were present both in monocot and dicot

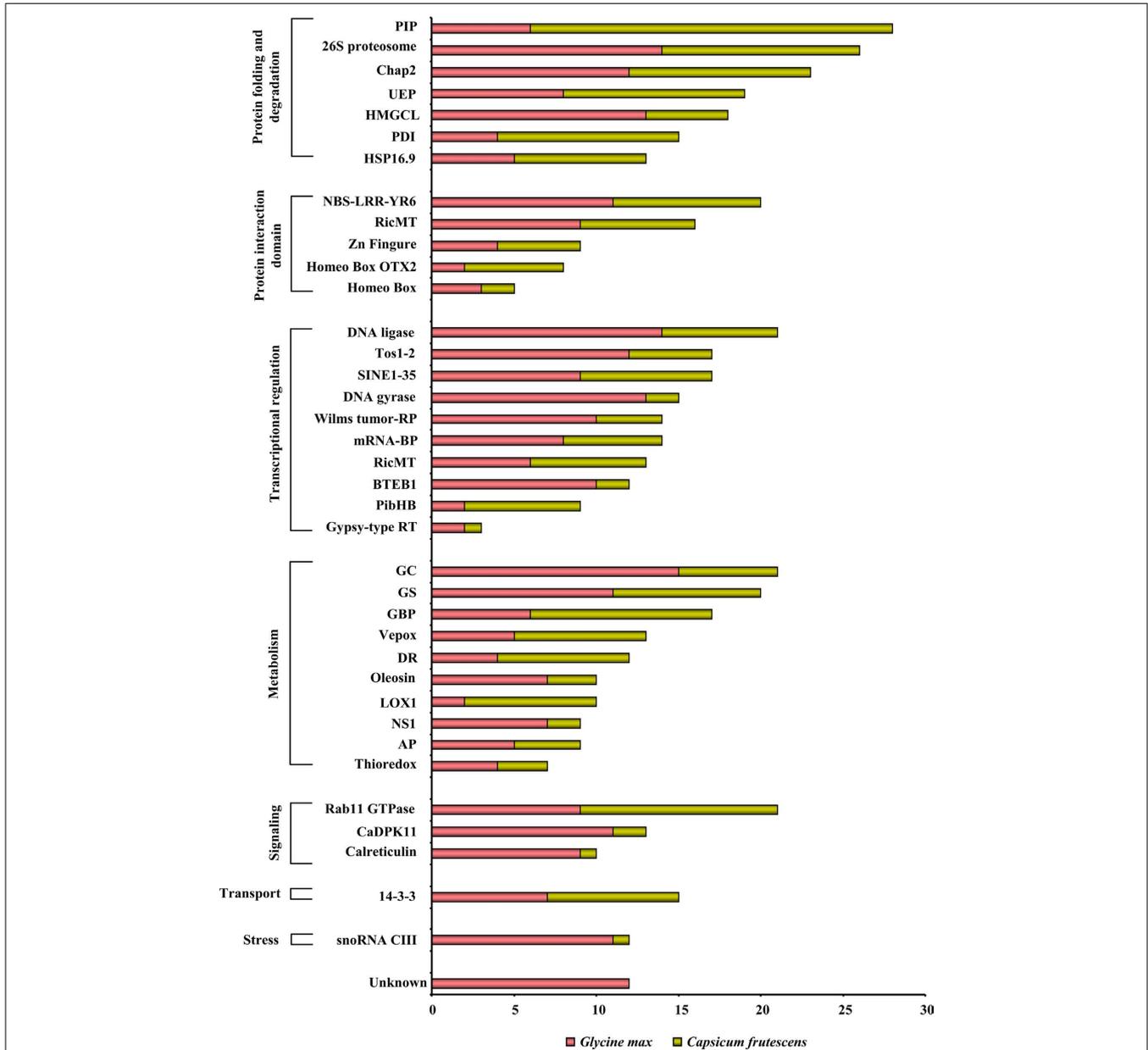


**FIGURE 3 | Comparative abiotic stress-responsive nuclear proteome: The functional classification of the identified proteins was according to the biological processes in which they are involved.** The length of the bars indicates the number of proteins present in a particular stress.

under varied stresses. Our analyses revealed the presence of monocot and dicot *cdc-2k*, *SEC31*, *TubA1* having specific protein sequences that clearly demonstrate the diversity of the identical NPs in two divisions of angiosperm. This may be attributed to the evolution of orthologs vs. paralogs.

Responses to various patho-stresses largely depend on the plant's capacity to modulate rapidly but specifically its proteome. External signals are translocated into the nucleus in a stress-type dependent manner to activate transcription factors, resulting in the increased expression of particular sets of defense-related genes. During evolution, mutual recognition between plants and

pathogens has resulted in development of fascinating variety of molecular strategies in the nucleus of the host against the invader. Some pathogens have been shown to directly activate transcription (Lev et al., 2005). It is now well accepted that modulation of chromatin configuration is an additional strategy employed by pathogen to subvert plant immune response (Ma et al., 2011). Nevertheless, plants also dispose an array of proteins in the nucleus that act as a scrutiny scheme to allow the early detection of an impending pathogen assault. We analyzed the nuclear proteomes of soybean and hot pepper in response to fungal (Cooper et al., 2011), and viral (Lee et al., 2006) stresses, respectively (Figure 4).



**FIGURE 4 | Comparative patho-stress responsive nuclear proteome: The functional classification of the identified proteins was according to the biological processes in which they are involved.** The length of the bars indicates the number of proteins present in a particular stress.

The widespread NPs identified in fungal and viral stresses belong to the category of protein folding and degradation. On the contrary, it was interesting enough to observe that there was not a single protein to be exclusive in case of either soybean-rust interaction or hot pepper–tobacco mosaic virus (TMV) interaction. During these host–pathogen interactions complex architecture of nucleus might respond differently against two different pathogens but using same set of NPs. Fungal stress and viral stress both might induce fundamental machinery of the nucleus to correctly target expressed proteins in a diverse but adaptation-related pathway thereby barricade the pathogens. However, NPs belonging to protein folding and degradation, transcription regulation, and metabolism categories toward patho-stress needs further consideration to understand the fungal-viral difference or specificity.

## CONCLUSION

Since its existence was first discovered almost 180 years ago, the nucleus has been a central focus of biological research. Initially it was assumed that nucleus is a static organelle. Progress over the years has gradually changed this view, and more recently, the importance of the NPs in chromatin organization, gene regulation, and signal transduction has become evident. In this study, cross-kingdom, cross-species as well as cross-condition comparisons of nuclear proteomes in vascular plants and animals illustrates the divergence in protein profiles within only a few social classes. *In silico* experimental analyses of the nuclear interior revealed a morphologically structured yet dynamic mix of NPs. Major nuclear events depend on the functional integrity of protein species and their timely interaction. Yet, unknown drivers of protein ensure that they are in the right place at the time when they are needed. Furthermore, the incessant unrest of proteins can be captured by the comparative nuclear proteome study under various regulatory events. As expected, the proteins involved in transcriptional regulation and chromatin remodeling were found to be the most predominant across all conditions. Nonetheless, a large number of proteins were unique or novel to each of the clades and under different stresses. It may be thought, the ubiquitously present protein classes are essential for sustenance, while the unique classes bring out the condition-specific special function. The differences in terms of protein pattern and protein function appear to encompass both genetic and physiological information. It may be speculated that the differential proteome is shaped by the cellular environment and the ecological niche of the corresponding organism. The divergence may arise due to codon bias, amino acid composition, and protein length. A much more comprehensive survey of the nuclear proteomes in several plants will ultimately draw a

more complete picture of the social class vs. protein diversity in this organelle.

## MARCHING AHEAD: NEXT FIVE YEARS

We are witnessing a significant but inadequate progress in understanding the nuclear proteomes of various crops of agricultural importance. Our understanding of nuclear composition, organization, and homeostasis has been greatly enhanced through targeted biochemical and genetic approaches. Unbiased “discovery” methods, such as proteomics, have only recently gained traction in the field of regulation biology. To date, a key word search using “Plant nuclear proteome” retrieves only 116 results in a pubmed search, emphasizing the need for in-depth study in the field. Although our knowledge of nuclear proteome and NPs has greatly increased, many open ended questions remain to be answered. It is to be noted that few thousands NPs identified in the nuclear proteomes have not been functionally characterized. Thus, for this new and emerging field, we predict that the potential for an accelerated pace of future discoveries in nuclear cell biology is tremendously high. The future scientific interest should center around the diverse roles NPs play in regulating cell division, growth, differentiation, aging, disease, and environmental perturbations. The comparative analysis of organism, clade-specific and stress-responsive plant nuclear proteomes revealed the presence of certain proteins that were unexpected, either in their abundance, form, number or else localization. These unexpected or non-canonical proteins suggest the constant remodeling of nuclear proteomes. The exact function and specificity of these candidates can only be comprehended once they are functionally characterized. Furthermore, role of PTMs on gene expression and NP-interactome dynamics remains as two important but challenging facets. Our future efforts will focus on the development and analysis of comparative nuclear proteomes toward an understanding of crop- and genotype-specific adaptation as an important amendment for the determination of protein networks influenced by the internal and external cues associated with the complex cellular, biochemical and physiological process that bring about phenome variation.

## ACKNOWLEDGMENTS

This research work was supported by grants from the Department of Biotechnology (DBT), Ministry of Science and Technology, Govt. of India (grant no. BT/PR10796/BRB/10/621/2008) and the National Institute of Plant Genome Research, India to S.C. K.N. is the recipient of pre-doctoral fellowship from Council of Scientific and Industrial Research (CSIR), Govt. of India. Authors thank Mr. Jasbeer Singh for illustrations and graphical representation in the manuscript.

## REFERENCES

- Abdalla, K. O., Baker, B., and Rafudeen, M. S. (2010). Proteomic analysis of nuclear proteins during dehydration of the resurrection plant *Xerophyta viscosa*. *Plant Growth Regul.* 62, 279–292.
- Abdalla, K. O., and Rafudeen, M. S. (2012). Analysis of the nuclear proteome of the resurrection plant *Xerophyta viscosa* in response to dehydration stress using iTRAQ with 2DLC and tandem mass spectrometry. *J. Proteomics* 18, 2361–2374.
- Abe, H., Urao, T., Seiki, M., and Shinozaki, Y. (2003). *Arabidopsis* AtMYC2 and AtMYB2 function as transcription regulator in abscisic acid signaling. *Plant Cell* 15, 68–73.
- Aki, T., and Yanagisawa, S. (2009). Application of rice nuclear proteome analysis to the identification of evolutionarily conserved and glucose-responsive nuclear proteins. *J. Proteome Res.* 8, 3912–3924.
- Bae, M. S., Cho, E. J., Choi, E. Y., and Park, O. K. (2003). Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *Plant J.* 36, 652–663.
- Barthelery, M., Salli, U., and Vrana, K. E. (2008). Enhanced nuclear proteomics. *Proteomics* 8, 1832–1838.
- Beven, A. F., Simpson, G. G., Brown, J. W., and Shaw, P. J. (1995). The organization of spliceosomal components in the nuclei of higher plants. *J. Cell Sci.* 108, 509–518.

- Bhushan, D., Pandey, A., Chatopadhyay, A., Choudhary, M. K., Chakraborty, S., Datta, A., et al. (2006). Extracellular matrix proteome of chickpea (*Cicer arietinum* L.) illustrates pathway abundance, novel protein functions and evolutionary perspective. *J. Proteome Res.* 5, 1711–1720.
- Bickmore, W. A., and Sutherland, H. G. E. (2002). Addressing protein localization within the nucleus. *EMBO J.* 21, 1248–1254.
- Boutillier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L. M., et al. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14, 1737–1749.
- Broers, J. L., Bronnenberg, N. M., Kuijpers, H. J., Schutte, B., Hutchison, C. J., and Ramaekers, F. C. (2002). Partial cleavage of A-type lamins concurs with their total disintegration from the nuclear lamina during apoptosis. *Eur. J. Cell Biol.* 81, 677–691.
- Brown, J. W. S., Shaw, P. J., Shaw, P., and Marshall, D. F. (2005). *Arabidopsis* nucleolar protein database (AtNoPDB). *Nucleic Acids Res.* 33, 633–636.
- Buhr, N., Carapito, C., Schaeffer, C., Kieffer, E., Van Dorsselaer, A., and Viville, S. (2008). Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. *Electrophoresis* 29, 2381–2390.
- Calikowski, T. T., Meulia, T., and Meier, I. (2003). A proteomic study of the *Arabidopsis* nuclear matrix. *J. Cell Biochem.* 90, 361–378.
- Casati, P. (2012). Recent advances in maize nuclear proteomic studies reveal histone modifications. *Front. Plant Sci.* 3:278. doi: 10.3389/fpls.2012.00278
- Casati, P., Campi, M., Chu, F., Suzuki, N., Maltby, D., Guan, S., et al. (2008). Histone acetylation and chromatin remodeling are required for UV-B-dependent transcriptional activation of regulated genes in maize. *Plant Cell* 20, 827–842.
- Caudron-Herger, M., and Rippe, K. (2012). Nuclear architecture by RNA. *Curr. Opin. Genet. Dev.* 22, 179–187.
- Cavalier-Smith, T. (1988). Origin of the cell nucleus. *BioEssays* 9, 72–78.
- Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int. J. Syst. Evol. Microbiol.* 52, 297–354.
- Cavalier-Smith, T. (2006). Cell evolution and earth history: stasis and revolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361, 969–1006.
- Cavalier-Smith, T. (2010). Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol. Direct* 5, 7–14.
- Cheung, A. Y., and Reddy, A. S. N. (2012). Nuclear architecture and dynamics: territories, nuclear bodies, and nucleocytoplasmic trafficking. *Plant Physiol.* 158, 23–25.
- Choudhary, M. K., Basu, D., Datta, A., Chakraborty, N., and Chakraborty, S. (2009). Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation and evolutionary perspective. *Mol. Cell. Proteomics* 8, 1579–1598.
- Constantinescu, D., Gray, H. L., Sammak, P. J., Schatten, G. P., and Csoka, A. B. (2006). Lamin A/C expression is a marker of mouse and human embryonic stem cell differentiation. *Stem Cells* 24, 177–185.
- Cooper, B., Campbell, K. B., Feng, J., Garrett, W. M., and Frederick, R. (2011). Nuclear proteomic changes linked to soybean rust resistance. *Mol. Biosys.* 3, 773–783.
- Cullen, Y. T., and Mansuy, I. M. (2010). Towards a better understanding of nuclear processes based on proteomics. *Amino Acids* 9, 1117–1130.
- Dahl, K. N., Ribeiro, A. J. S., and Lammerding, J. (2008). Nuclear shape, mechanics, and mechanotransduction. *Circ. Res.* 102, 1307–1318.
- De Souza, C. P., and Osmani, S. A. (2009). Double duty for nuclear proteins – the price of more open forms of mitosis. *Trends Genet.* 25, 545–554.
- Dundr, M., and Misteli, T. (2001). Functional architecture in the cell nucleus. *Biochem. J.* 356, 297–310.
- Erhardt, M., Adamska, I., and Luiz Franco, O. (2010). Plant nuclear proteomics – inside the cell maestro. *FEBS J.* 16, 3295–3307.
- Feng, B., Yang, S., Du, H., Hou, X., Zhang, J., Liu, H., et al. (2012). Molecular characterization and functional analysis of plant WRKY genes. *African J. Biotechnol.* 11, 13606–13613.
- Ferreira, B. S., Branco, A. T., deOliveira, M. A., Pereira, M. G., and de Souza Filho, G. A. (2006). Methodological improvements on extraction of nuclear proteins and its preliminary analysis during the maize (*Zea mays* L.) endosperm development. *Protein Pept. Lett.* 13, 981–984.
- Franklin, S., Zhang, M. J., Chen, H., Paulsson, A. K., Jordan, S. S. M., Li, Y., et al. (2011). Specialized compartments of cardiac nuclei exhibit distinct proteomic anatomy. *Proteomics* 10, 1074–1015.
- Fritz-Laylin, L. K., Prochnik, S. E., Ginger, M. L., Dacks, J. B., Carpenter, M. L., Field, M. C., et al. (2010). The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell* 140, 631–642.
- Gao, G., Zhong, Y., Guo, A., Zhu, Q., Tang, W., Zheng, W., et al. (2006). DRTF: a database of rice transcription factors. *Bioinformatics* 15, 1286–1287.
- Gauci, S., Veenhoff, L. M., Heck, A. J. R., and Krijgsveld, J. (2009). Orthogonal separation techniques for the characterization of the yeast nuclear proteome. *J. Proteome Res.* 8, 3451–3463.
- Gonzalez-Melendi, P., Beven, A., Boudonck, K., Abranches, R., Wells, B., Dolan, L., et al. (2000). The nucleus: a highly organized but dynamic structure. *J. Microsc.* 198, 199–207.
- Gorski, S., and Misteli, T. (2005). Systems biology in the cell nucleus. *J. Cell Sci.* 118, 4083–4092.
- Guo, A., He, K., Liu, D., Bai, S., Gu, X., Wei, L., et al. (2005). DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* 15, 2568–2569.
- Guo, A. Y., Chen, X., Gao, G., Zhang, H., Zhu, Q. H., Liu, X. C., et al. (2008). PlantTFDB: a comprehensive plant transcription factor database. *Nucleic Acids Res.* 36, 966–969.
- Henrich, S., Cordwell, S. J., Crossett, B., Baker, M. S., and Christopherson, R. I. (2007). The nuclear proteome and DNA-binding fraction of human Raji lymphoma cells. *Biochim. Biophys. Acta* 1774, 413–432.
- Hetzer, M. W., Walther, T. C., and Mattaj, I. W. (2005). Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annu. Rev. Cell Dev. Biol.* 21, 347–380.
- Holmes-Davis, R., and Comai, L. (1998). Nuclear matrix attachment regions and plant gene expression. *Trends Plant Sci.* 3, 91–97.
- Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, T. T., Konagaya, A., et al. (2005). RARTE: database and tools for complete sets of *Arabidopsis* transcription factors. *DNA Res.* 12, 247–256.
- Jiang, G., Huang, A. H., Cai, Y., Tanase, M., and Sheetz, M. P. (2006). Rigidity sensing at the leading edge through alpha-beta3 integrins and RPTPalpa. *Biophys. J.* 90, 1804–1809.
- Khan, M. K. K., and Komatsu, S. (2004). Rice proteomics: recent developments and analysis of nuclear proteins. *Phytochemistry* 65, 1671–1681.
- Lee, B. J., Kwon, S. J., Kim, S. K., Kim, K. J., Park, C. J., Kim, Y. J., et al. (2006). Functional study of hot pepper 26S proteasome subunit RPN7 induced by tobacco mosaic virus from nuclear proteome analysis. *Biochem. Biophys. Res. Comm.* 351, 405–411.
- Lev, S., Hadar, R., Amedeo, P., Baker, S. E., Yoder, O. C., and Horowitz, B. A. (2005). Activation of an AP1-Like transcription factor of the maize pathogen *Cochliobolus heterostrophus* in response to oxidative stress and plant signals. *Eukaryot. Cell* 4, 443–454.
- Li, G., Nallamilli, B. R., Tan, F., and Peng, Z. (2008). Removal of high abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 29, 604–617.
- Liao, L., McClatchy, D. B., and Yates, J. R. (2009). Shotgun proteomics in neuroscience. *Neuron* 63, 12–26.
- Ma, K. W., Flores, C., and Ma, W. (2011). Chromatin configuration as a battlefield in plant-bacteria interaction. *Plant Physiol.* 157, 535–543.
- Mans, B. J., Anantharaman, V., Aravind, L., and Koonin, E. V. (2004). Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637.
- Mayrose, M., Ekengren, S. K., Melech-Bonfil, S., Martin, G. B., and Sessa, G. (2006). A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. *Mol. Plant Pathol.* 7, 593–604.
- McClatchy, D. B., Liao, L., Park, S. K., Xu, T., Lu, B., and Yates, J. R. (2011). Differential proteomic analysis of mammalian tissues using SILAM. *PLoS ONE* 6:e16039. doi: 10.1371/journal.pone.0016039
- Meagher, R. B., McKinney, E. C., and Kandasamy, M. K. (1999). Isovariant dynamics expand and buffer the responses of complex systems: the diverse plant actin gene family. *Plant Cell* 11, 995–1005.
- Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291, 843–847.
- Misteli, T. (2005). Concepts in nuclear architecture. *BioEssays* 27, 477–487.
- Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y., and Kurata, N. (2005). Functional isolation of novel nuclear proteins showing a variety of subnuclear localizations. *Plant Cell* 17, 389–403.
- Narula, K., Elagamey, E., Datta, A., Chakraborty, N., and Chakraborty, S. (2012). “Comparative analyses of extracellular matrix proteome:

- an under-explored area in plant research," in *Crop Plants*, ed. A. Goyal (Janeza Tradine, Croatia: InTech), 145–166.
- Pajerowski, J. D., Dahl, K. N., Zhong, F. L., Sammak, P. J., and Discher, D. E. (2007). Physical plasticity of the nucleus in stem cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15619–15624.
- Palaniswamy, S. K., James, S., Sun, H., Lamb, R. S., Davuluri, R. V., and Grotewold, E. (2006). AGRIS and AtRegNet. A platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol.* 140, 818–829.
- Pandey, A., Chakraborty, S., Datta, A., and Chakraborty, N. (2008). Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* 7, 88–107.
- Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A., Chakraborty, S., Datta, A., et al. (2006). The nuclear proteome of chickpea (*Cicer arietinum* L.) reveals predicted and unexpected proteins. *J. Proteome Res.* 5, 3301–3311.
- Pederson, T. (2000). Diffusional protein transport within the nucleus: a message in the medium. *Nat. Cell Biol.* 2, E73–E74.
- Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., et al. (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol. Biol. Cell* 16, 260–269.
- Pennisi, E. (2004). The birth of the nucleus. *Science* 305, 766–768.
- Perez-Rodriguez, P., Riano-Pachon, D. M., Correa, L. G., Rensing, S. A., Kersten, B., and Roeber, M. B. (2010). PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 38, 822–827.
- Renouard, S., Cyrielle, C., Lopez, T., Lamblin, F., Lainé, E., and Hano, C. (2012). Isolation of nuclear proteins from flax (*Linum usitatissimum* L.) seed coats for gene expression regulation studies. *BMC Res. Notes* 5:15. doi: 10.1186/1756-0500-5-15
- Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Küster, H., Larré, C., et al. (2008). Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 56, 398–410.
- Repetto, O., Rogniaux, H., Larré, C., Thompson, R., and Gallardo, K. (2012). The seed nuclear proteome. *Front Plant Sci.* 3:289. doi: 10.3389/fpls.2012.00289
- Rivas, S. (2012). Nuclear dynamics during plant innate immunity. *Plant Physiol.* 158, 87–94.
- Roix, J., and Misteli, T. (2002). Genomes, proteomes, and dynamic networks in the cell nucleus. *Histochem. Cell Biol.* 118, 105–116.
- Romeuf, I., Tessier, D., Dardevet, M., Branlard, G., Charmet, G., and Ravel, C. (2010). wDBTF: an integrated database resource for studying wheat transcription factor families. *BMC Genomics* 11:185. doi: 10.1186/1471-2164-11-185
- Rushton, P. J., Bokowiec, M. T., Lauderman, T. W., Brannock, J. F., Chen, X., and Timko, M. P. (2008). TOBFAC: the database of tobacco transcription factors. *BMC Bioinformatics* 9:53. doi: 10.1186/1471-2105-9-53
- Salzano, A. M., Paron, I., Pines, A., Bachi, A., Talamo, F., Bivi, N., et al. (2006). Differential proteomic analysis of nuclear extracts from thyroid cell lines. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 833, 41–50.
- Schirmer, E. C., and Gerace, L. (2005). The nuclear membrane proteome: extending the envelope. *Trends Biochem. Sci.* 30, 551–558.
- Semple, C. A., Riken, G. E. R., and Members, G. S. L. (2003). The comparative proteomics of ubiquitination in mouse. *Genome Res.* 13, 1389–1394.
- Skovgaard, M., Jensen, L. J., Brunak, S., Ussery, D., and Krogh, A. (2001). On the total number of genes and their length distribution in complete microbial genomes. *Trends Genet.* 17, 425–428.
- Spector, D. L. (2003). The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* 72, 573–608.
- Strouboulis, J., and Wolffe, A. P. (1996). Functional compartmentalization of the nucleus. *J. Cell Sci.* 109, 1991–2000.
- Su, P. A., and Li, H. M. (2008). *Arabidopsis* stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant Physiol.* 146, 1231–1241.
- Takizawa, T., and Meshorer, E. (2008). Chromatin and nuclear architecture in the nervous system. *Trends Neurosci.* 31, 345–355.
- Tan, F., Li, G., Chitteti, B. R., and Peng, Z. (2007). Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*). *Proteomics* 7, 4511–4527.
- Tan, F., Zhang, K., Mujahid, H., Verma, D. P., and Peng, Z. (2010). Differential histone modification and protein expression associated with cell wall removal and regeneration in rice (*Oryza sativa*). *J. Proteome Res.* 10, 551–563.
- Truncaite, L., Zajanckauskaite, A., Arlauskas, A., and Nivinskas, R. (2006). Transcription and RNA processing during expression of genes preceding DNA ligase gene 30 in T4-related bacteriophages. *Virology* 344, 378–390.
- van Driel, R., and Fransz, P. (2004). Nuclear architecture and genome functioning in plants and animals: what can we learn from both? *Exp. Cell Res* 296, 86–90.
- Varma, P., and Mishra, R. K. (2011). Dynamics of nuclear matrix proteome during embryonic development in *Drosophila melanogaster*. *J. Biosci.* 36, 439–459.
- Vergnes, L., Peterfy, M., Bergo, M. O., Young, S. G., and Reue, K. (2004). Lamin B1 is required for mouse development and nuclear integrity. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10428–10433.
- Wang, Z., Libault, M., Joshi, T., Valliyodan, B., Nguyen, H. T., Xu, D., et al. (2010). SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biol.* 10:14. doi: 10.1186/1471-2229-10-14
- Wilson, K. L., and Dawson, S. C. (2011). Functional evolution of nuclear structure. *J. Cell Biol.* 195, 171–181.
- Winck, F. V., Riano-Pacho, D. M., Sommer, F., Rupperecht, J., and Roeber, B. M. (2011). The nuclear proteome of the green alga *Chlamydomonas reinhardtii*. *Proteomics* 12, 95–100.
- Yan, W., Hwang, D., and Aebersold, R. (2008). Quantitative proteomic analysis to profile dynamic changes in the spatial distribution of cellular proteins. *Methods Mol. Biol.* 432, 389–401.
- Yates, J. R., Gilchrist, A., Howell, K. E., and Bergeron, J. J. (2005). Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell Biol.* 6, 702–714.
- Yu, Y., Xu, W., Wang, S., Xu, Y., Li, H., Wang, Y., et al. (2011). VpRFP1, a novel C4C4-type RING finger protein gene from Chinese wild *Vitis pseudoreticulata*, functions as a transcriptional activator in defence response of grapevine. *J. Exp. Bot.* 62, 671–682.
- Zhu, Q. H., Guo, A. Y., Gao, G., Zhong, Y. F., Xu, M., Huang, M., and Luo, J. (2007). DPTF: a database of poplar transcription factors. *Bioinformatics* 15, 1307–1308.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2013; paper pending published: 01 February 2013; accepted: 30 March 2013; published online: 26 April 2013.

Citation: Narula K, Datta A, Chakraborty N and Chakraborty S (2013) Comparative analyses of nuclear proteome: extending its function. *Front. Plant Sci.* 4:100. doi: 10.3389/fpls.2013.00100

This article was submitted to *Frontiers in Plant Proteomics*, a specialty of *Frontiers in Plant Science*.

Copyright © 2013 Narula, Datta, Chakraborty and Chakraborty. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.