Dynamics of nuclear matrix proteome during embryonic development in *Drosophila melanogaster*

Parul Varma and Rakesh K Mishra*
Center for Cellular and Molecular Biology (CSIR), Hyderabad 500 007, India

**Corresponding author (Fax, +91-40-27160591; Email, mishra@ccmb.res.in)*

Embryonic development is a complex and dynamic process that involves spatiotemporal expression of genes in a highly coordinated manner. Multiple levels of nuclear architecture maintain the fidelity of gene expression programme. One of the components of nuclear architecture, which is believed to play an important role in regulation of gene expression, is the nuclear matrix (NuMat). Many studies over the past few years have tried to analyse the components of this non-chromatin scaffolding of the nucleus and have provided evidences of its structural and functional complexity. However, the relationship of NuMat with the process of embryonic development still remains poorly understood. Here, we report a comparative analysis of the NuMat proteomes of early and late stage *Drosophila melanogaster* embryos and show that 65% of the NuMat proteome is dynamic during development. Our study establishes links between the dynamics of nuclear architecture and embryonic development and provides tools to further understand the process such as cellular differentiation in the context of higher-order nuclear organization.

[Varma P and Mishra RK 2011 Dynamics of nuclear matrix proteome during embryonic development in *Drosophila melanogaster*. *J. Biosci.* **36** 439–459] **DOI** 10.1007/s12038-011-9081-6

1. Introduction

Multicellular organisms arise by a relatively slow and complex process of progressive changes called development. In nearly all cases, development begins with a single cell – the fertilized egg or zygote, which divides mitotically to produce all the cells of the body. Molecularly, this process is not only a highly dynamic orchestrated event of switching on and off the transcription of various genes, depending upon the stage of development, but also a synergism between the processes rendering the *cis*-elements accessible or refractory to various *trans*-factors and the positioning of genetic material in the nuclear space, which is compartmentalized into distinct functional and structural domains (Jackson 2003). While the functional domains include replication foci, transcription foci, Cajal bodies, promyelocytic leukaemia protein (PML) bodies, nuclear speckles, etc., the structural domains include heterochromatin, nucleolus, nuclear lamina and nuclear territories (Lanctot *et al*. 2007). The different levels of non-random packaging and folding of the genome itself also results in various chromatin territories and domains (Cremer *et al*. 2004). All these nuclear parameters together contribute to the higher-order nuclear architecture and its role in regulation of gene expression. However, the establishment and maintenance of these structural and functional domains and, consequently, the interplay between the nuclear architecture and function during various biological processes involving development and differentiation still remains poorly understood. It is hypothesized that this interplay is draped over the non-chromatin structure of the nucleus called nuclear matrix (NuMat) (Berezney and Wei 1998).

NuMat, hidden beneath the large mass of chromatin, is defined as a biochemical fraction that is isolated after sequential extractions with non-ionic detergents, nucleases and high-salt buffers (Berezney and Coffey 1977). The structural entities of this nuclear fraction include residual elements of the nuclear envelope, also known as the pore complex-lamina, residual nucleoli and a granular and fibrous internal meshwork, which extends throughout the interior of the nucleus (Capco *et al*. 1982). Molecularly, NuMat consists of DNA (MARs /SARs), RNA and proteins

Keywords. Embryonic development; nuclear architecture; nuclear matrix; NuMat proteome

(Jackson and Cook 1988). MARs are the 'matrix attachment regions' and SARs are the 'scaffold attachment regions' of the DNA. Many earlier studies, in an attempt to characterize the constituents of this structural framework of the nucleus, have been successful in identifying a small but diverse group of proteins, reflecting a link between nuclear architecture and its involvement in a variety of nuclear processes such as replication (Anachkova *et al*. 2005), transcription (S'Iakste and S'Iakste 2001), hnRNA processing (Zeitlin *et al*. 1987), etc.

More recently, major advances in the field of proteomics have unveiled the otherwise poorly understood protein composition of this nuclear structural framework from various kinds of cells and tissues in a variety of biological and pathological conditions (Albrethsen *et al*. 2009). In fact, in a recent study, an extensive analysis of NuMat proteome composition has reflected the rich variety of functional categories of proteins associated with NuMat, providing further implications of its role in the diverse nuclear processes. These proteins included structural proteins, heat shock proteins, chaperones, DNA binding and mitotic scaffolding proteins, chromatin remodeling proteins, transcription- and translation-associated proteins and also ribosomal proteins (Kallappagoudar *et al*. 2010). Taken together, these studies have further strengthened the concept of NuMat being associated with processes such as embryonic development and differentiation and hence the interplay between the dynamics of nuclear architecture and regulation of gene expression during these key molecular events.

In the present study, we carried out a comparative analysis between the NuMat proteome of two different developmental stages (early and late) of *Drosophila melanogaster* embryos. NuMat preparations from these embryos were identified by LC-MS/MS. While our earlier report shed light on the changing pattern of the NuMat proteome during development (Kallappagoudar *et al*. 2010), this study shows that NuMat proteome is indeed dynamic with remarkable variation between the two developmental stages. Thus, this analysis establishes a link between the dynamics of both embryonic development as well as nuclear architecture.

2. Materials and methods

2.1 *Embryo collection*

Canton S flies were used for all the experiments. Plates containing solidified fly food base streaked with yeast paste were placed inside the large cages with these wild-type flies. For early stage (0–2 h) embryo collection, the plates were removed from the cages after 2 h and embryos were collected. For late stage (14–16 h) embryo collection, plates were removed from the cages after 2 h but kept at 25°C for another 14 h before collection. This staging made the embryos 14–16 h old.

Once the embryos were collected, they were dechorionated with 50% Chlorex (sodium hypochlorite 4–6%) for approximately 5 min to remove chorion. The dechorionated embryos were taken in a mesh and washed thoroughly under running water to remove Chlorex completely. These embryos were further processed for NuMat preparations.

2.2 *Isolation of NuMat*

Dechorionated embryos were homogenized in pre-chilled 0.25 M sucrose in nuclear isolation buffer (NIB, 15 mM Tris pH 7.4, 40 mM KCl, 1 mM EDTA, 0.1 mM EGTA, 0.25 mM Spermidine, 0.1 mM Spermine, 0.1 mM PMSF and 0.1 µM Aprotinin) in a motor-driven glass homogenizer (10 strokes) and filtered through two layers of Mira cloth (Calbiochem). The filtrate was spun at 600*g* at 4°C for 2 min to remove the crud. The resulting supernatant was mixed gently with an equal volume of 1.8 M sucrose in NIB and centrifuged at 6000*g* at 4°C to isolate the nuclear pellet. The pellet obtained was left in NIB containing 0.25 M sucrose for 5 min at 4°C, and then washed twice by re-suspending it completely in NIB with 0.25 M sucrose and spinning at 3000*g* for 10 min at 4°C. The washed nuclear pellet was incubated in digestion buffer (20 mM Tris pH 7.4, 20 mM KCl, 70 mM NaCl, 10 mM $MgCl_2$, 0.125 mM Spermidine, 0.05 mM Spermine, 0.1 mM PMSF, 0.1 µM Aprotinin, 0.5% Triton X-100 and 40 U/µl DNase I) at 4°C for 1 h. Digestion was followed by first extraction for 5 min with 0.4 M NaCl and then for another 5 min by increasing the salt concentration to 2 M in extraction buffer (10 mM Hepes pH 7.5, 4 mM KCl, 4 mM EDTA, 0.5 mM Spermidine, 0.1 mM PMSF, 0.1 µM Aprotinin and 0.5% Triton X-100) at room temperature. The final pellet obtained was washed twice with wash buffer (5 mM Tris pH 7.5, 20 mM KCl, 1 mM EDTA, 0.25 mM Spermidine, 0.1 mM Spermine, 0.1 mM PMSF and 0.1 µM Aprotinin) by spinning at 3000*g* at 25°C. The final pellet was stored at −70°C or used for further analysis.

2.3 *Gel electrophoresis and in gel digestion for LC-MS/MS*

For each LC-MS/MS experiment, 200 µg of nuclear matrix preparations, from 0–2 h and 14–16 h old embryos were fractionated on a 12% SDS-PAGE gel. The gels were stained with Coomassie brilliant blue (CBB, R250) for 1 h, destained and washed thoroughly with MilliQ H_2O. Each gel lane was sliced into equal number of pieces. These gel slices were chopped finely, taken into clean microfuge tubes and washed three times for 30 min each in 500 µl of 50% Acetonitrile (ACN), 25 mM ammonium bicarbonate (ABC), pH 8.0, by gentle agitation at room temperature to remove

excess CBB stain. The gel pieces were given a final wash for 15 min with 500 µl of 50% ACN and 10 mM ABC to remove excess salt. After washings, these gel slices were soaked in 500 µl of 100% ACN for 5 min to dehydrate the gels. ACN was completely removed after 5 min and the dehydrated gel slices were further vacuum dried for 30 min. These dried gels were rehydrated and trypsinized with 30–40 µl cold Trypsin (Promega) solution (10 µg/ml in 25 mM ABC, pH 8.0) and incubated at 37°C for 16 h. The tryptic peptides were extracted by soaking the gel slices in 50 µl of 50% ACN and 5% trifluoroacetic acid (TFA) for 1 h with gentle agitation at room temperature. The supernatant was collected and transferred to a new clean microfuge tube. The gels were extracted again with another aliquot of 50% ACN and 5% TFA under the same conditions. The two tryptic extracts were pooled and vacuum-dried to complete dryness for 1 h. These dried extracts were stored at −80°C. Before loading, the samples were reconstituted in 12 µl of 5% ACN and 0.1% formic acid and spun at 10000 rpm for 10 min to remove any insoluble particles.

2.4 *LC-MS/MS analysis*

All LC-MS/MS analysis were carried out on an ESI-mass spectrometer with linear ion trap mass analyser (LTQ-IT, Thermo Fischer, Waltham, MA, USA), equipped with Finnigan Surveyor MS Pump Plus. 10 µl of the sample was loaded with constant flow of 2 µl/min onto a reverse-phase Micro LC column (Bio Basic $C_{18,}$ Thermo Fischer). Peptides were eluted on a gradient of 90 min for each gel slice starting with 5% ACN, 95% H_2 O for the first 10 min. ACN gradient was increased from 5% to 95%, while H_2 O was reduced from 95% to 5% over the next 65 min. The 95% ACN, 5% H_2 O gradient was maintained for another 5 min and then changed to 5% ACN, 95% H_2 O over the next 10 min. Chromatographically separated peptides were sprayed through 20 cm needle emitter and the mass spectrometer was operated in the data-dependent mode to acquire MS and MS/MS spectra switching automatically between MS and MS/MS modes. One full MS scan from 300 to 2000 m/z followed by 7 data-dependent MS/MS scans was recorded. The electrospray voltage was set at 4.5 kV, and the capillary temperature at 200°C. The peptides were fragmented using CID with normalized collision energy of 35%. The top 7 peptide precursor ions were selected for MS/MS analysis. The raw files were compiled and subjected to bioinformatics analysis.

2.5 *Bioinformatics analysis*

The mass spectra obtained were searched against the protein sequences of the *D. melanogaster* assembly 5.2 obtained from NCBI (20, 513 proteins) using SEQUEST (Eng *et al*. 1994) algorithm incorporated in the BioWorks Browser (Version 3.2 EF2, Thermoelectron Corp, Waltham, MA, USA). Enzyme specificity was set to full trypsin digestion with the allowance of only one missed cleavage. Methionine oxidation was allowed as a variable modification. Precursor ion tolerance was set to 1 amu and fragment ion tolerance was set to 0.35 amu. Peptide identifications were accepted if they passed the following filter criteria: ΔCN of 0.100; Rsp of 5; Xcorr *versus* charge values of 1.90 (+1 charge), 2.20 (+2 charge) and 3.30 (+3 charge); and protein probability of 0.001. An in-house program was used to remove redundant peptides and data compilation (Kallappagoudar *et al*. 2010)

2.6 *Statistical analysis*

The percentage of unique proteins/functional group/stage was calculated using the following formula:

$$\%\ \text{Unique Proteins}\ (UP)\{\mathrm{X},\mathrm{Y}\} = (A/N) \times 100,$$

where X represents functional group,Y represents developmental stage, A = number of unique proteins in functional group X in stage Y and N = total number of unique proteins in stage Y.

The percentage of common proteins/functional group was calculated using the following formula:

$$\%\ \text{Common Proteins}\,(CP)\{\mathrm{X}\} = (B/N) \times 100,$$

where X represents functional group, B = common number of proteins present in all the stages for functional group X and N = total number of common proteins.

Complexity pattern of each class was calculated as follows:

$$\text{Complexity shift}\,CS\{\mathrm{X}\} = C - D,$$

where X represents functional group, C = number of proteins unique to early stage (0–2 h) NuMat proteome in functional group X, D = Number of proteins unique to late stage (14–16 h) NuMat proteome in functional group X, $CS>0$ represents 'early complexity' class and $CS<0$ represents 'late complexity' class.

The % dynamics of each functional category as well as the whole proteome was calculated using the following formula:

$$\%\ \text{Dynamics}\ (\mathrm{X}) = \{(a+b)/(a+b+c)\} \times 100,$$

where X represents functional group, a = number of proteins unique to early stage (0–2 h) NuMat proteome in functional group X, b = number of proteins unique to late stage (14–16 h) NuMat proteome in functional group X and c = number of proteins common between the NuMat proteomes of both the stages in functional group X.

The degree of dynamics of each functional class was assigned based on the formula:

$$\text{Range}\,(R) = (E - F),$$

where E = highest value of % dynamics observed and F = lowest value of % dynamics observed. By dividing the range (R) in three equal parts, the class intervals were defined as less dynamic, moderately dynamic and highly dynamic representing an increase in the degree of dynamics successively.

3. Results

3.1 *LC-MS/MS profiling of the NuMat proteome of early and late stage D. melanogaster embryos*

NuMat is the nuclease and salt-resistant fraction of the nuclei enriched in high-molecular-weight proteins (Kallappagoudar *et al*. 2010). Figure 1A represents a standard NuMat proteome profile from 0- to 16-h-old *Drosophila* embryos, while figure 1B represents the NuMat preparations from early (0–2 h) and late (14–16 h) stage embryos. An earlier report using two-dimensional difference gel electrophoresis (DIGE) had shown that NuMat profiles of early and late stage *Drosophila* embryos are remarkably variant (Kallappagoudar *et al*. 2010). This prompted us to identify the NuMat proteomes of these two developmental stages to evaluate the NuMat dynamics during the process of embryonic development (table 1).

We used liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) for differential proteome profiling as described under 'materials and methods'. Biological replicates were used for the study, and it was found that most of the proteins were reproducibly detected in all the experiments. However, the protein lists presented here are inclusive of all the replicates (tables 2, 3 and 4). The total number of proteins identified in early and late stage NuMat was 267 and 295, respectively. Out of 267 proteins in early stage NuMat, 98 proteins were identified by two or more than two unique peptides while 169 proteins accounted for single peptide hits. Similarly, out of 296 proteins identified in late stage NuMat, 112 were identified by two or more than two unique peptides while the remaining 183 was only single peptide hits. We think that a large proportion of these proteins are bonafide NuMat constituents as there are many proteins for which the single peptide is the only peptide that can be detected. This is often due to the less abundance or the smaller size of the protein. In our earlier study, boundary-element-associated factor (BEAF), a 32 kDa protein, which is one of the known NuMat constituent, was identified by only single peptide hit (Kallappagoudar *et al*. 2010; Pathak *et al*. 2007). Our data also shows here that less dynamic part of NuMat proteome,

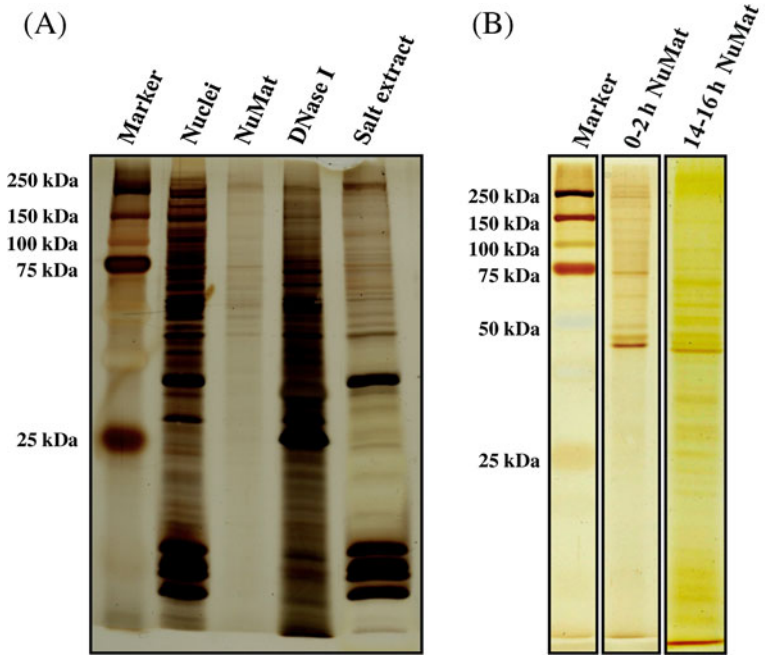

Figure 1. NuMat proteome profiles. (**A**) Silver-stained PAGE with extracts at different step in NuMat preparation from 0- to 16-h-old *D. melanogaster* embryos. NuMat seen in the salt-resistant fraction is enriched in higher-molecular-weight proteins. (**B**) A comparison between the silver-stained profiles of NuMat preparations from early (0–2 h) and late stage (14–16 h) embryos.

Table 1. Comparative analysis of the classes of proteins associated with 0–2 h and 14–16 h NuMat proteome

S.No.	Functional class	Relative complexity*	Dynamics (%)†	Remarks
1	Nuclear membrane/Nuclear pore	7 9 4	55	Early complexity Moderately dynamic
2	Structural proteins	3 11 18	68	Late complexity Highly dynamic
3	Heat shock proteins/Chaperones	3 16 1	20	Early complexity Least dynamic
4	Enzymes/Metabolic process	23 11 8	74	Early complexity Highly dynamic
5	DNA binding/Replication	15 6 7	79	Early complexity Highly dynamic
6	Cell cycle/Cell proliferation	10 2 4	88	Early complexity Highly dynamic
7	Chromatin remodeling	5 12 6	48	Late complexity Moderately dynamic
8	Transcription	2 2 14	89	Late complexity Most dynamic
9	RNA binding/mRNA processing	8 5 4	71	Early complexity Highly dynamic
10	Translation	8 7 17	78	Late complexity Highly dynamic
11	Transport	4 1 2	86	Early complexity Highly dynamic
12	Signal transduction	3 4 4	64	Late complexity Moderately dynamic
13	Mitochondrial	3 2 2	71	Early complexity Highly dynamic
14	Ribosomal	4 35 21	42	Late complexity Less dynamic
15	Others	22 24 36	71	Most complex Highly dynamic

*The bar graphs representing the complexity pattern of respective functional groups are drawn to the scale. The proteins unique to 0–2 h, unique to 14–16 h and common to both the stages in the respective functional groups are shown in green, blue and red color, respectively. The numbers preceding the bars depict the absolute number of proteins (not percentage) in each functional group in the respective stage.

† Indicates % of proteins in the corresponding class that vary between the age group embryos.

Early and late complexity indicates that the complexity of the NuMat proteome in reference to the number of proteins in a particular functional group is more in the early stage or late stage of the development, respectively.

Table 2. Proteins unique to 0–2 h NuMat proteome

S.No.	Accession Number	Protein Name	Homologues*						Total peptide hits
		I. Nuclear membrane/Nuclear Pore							
1	NP_650682.1	cadmus	-	-	-	I	V	H	1
2	NP_651080.2	CG6958	-	-	-	I	V	H	1
3	NP_477496.1	Female sterile (2) Ketel	F	P	W	I	-	H	2
4	NP_651361.2	Nup358	F	P	-	-	V	H	4
5	NP_650404.1	CG7262	F	P	W	I	V	H	2
6	NP_477041.1	Pendulin	-	-	-	I	V	H	1
7	NP_650068.1	Ranbp9	-	P	-	I	V	H	1
		II. Structural proteins							
1	NP_523929.2	Dynein heavy chain 64 C	-	-	W	I	V	H	1
2	NP_724495.1	Eb1	-	P	-	I	V	H	2
3	NP_727910.1	shibire	-	-	W	I	V	H	2
		III. Heat shock proteins/Chaperones							
1	NP_523936.2	DnaJ-like-1	-	P	-	I	V	H	1
2	NP_725084.2	ERp60	-	P	W	I	V	H	1
3	NP_477439.2	Trap1	-	-	W	I	V	H	2
		IV. Enzymes/Metabolic process proteins							
1	NP_620470.1	26–29 kDa-proteinase	-	-	-	I	V	-	2
2	NP_572826.1	ade5	-	-	W	I	V	H	2
3	NP_001097556.1	Arginine kinase	-	-	-	-	-	-	1
4	NP_609336.3	CG4747	-	P	-	I	V	H	3
5	NP_648484.1	CG6084	F	P	W	I	V	H	1
6	NP_729528.2	CG6767	-	P	-	I	-	H	1
7	NP_729517.1	CG8336	F	P	-	I	V	H	1
8	NP_608375.1	CG9577	-	P	W	I	V	H	1
9	NP_609479.1	Deoxyuridine triphosphatase	F	P	W	I	-	H	1
10	NP_524452.4	Diphthamide methyltransferase	F	P	W	I	V	H	1
11	NP_001036476.1	Down syndrome cell adhesion molecule	-	-	-	I	V	H	1
12	NP_524326.1	Glutathione S transferase D1	-	-	-	-	-	-	2
13	NP_650181.1	Glutathione S transferase D9	-	-	-	-	-	-	1
14	NP_725653.1	Glutathione S transferase S1	-	-	W	I	-	H	1
15	NP_542445.1	Glyceraldehyde 3 phosphate dehydrogenase 2	F	P	W	I	V	H	2
16	NP_649466.1	growl	-	-	-	-	-	-	1
17	NP_523969.1	*N*-myristoyl transferase	-	P	-	I	V	H	1
18	NP_476676.1	Phosphoglycerate kinase	F	-	W	I	V	H	2
19	NP_524546.2	Phosphoglyceromutase	-	-	-	I	V	H	2
20	NP_523532.1	Proteasome 35 kDa subunit	F	P	W	I	V	H	1
21	NP_524837.2	Proteasome alpha6 subunit	F	P	W	I	V	H	1
22	NP_477378.1	Thiolase	-	-	W	I	V	H	2
23	NP_477310.2	Ubiquitin activating enzyme 1	F	P	W	I	V	H	1
		V. DNA binding/Replication							
1	NP_648247.1	CG5971	F	-	W	I	V	H	1
2	NP_650715.3	CG7670	-	-	-	-	-	-	3
3	NP_611019.2	CG8092	-	-	-	-	-	-	1
4	NP_573245.1	CG8142	F	P	-	I	V	H	2
5	NP_476904.1	latheo	-	P	-	I	V	H	1
6	NP_477303.1	Origin recognition complex subunit 1	-	-	-	I	V	H	3

Table 2. (continued)

S.No.	Accession Number	Protein Name	Homologues*						Total peptide hits
7	NP_731873.1	Origin recognition complex subunit 2	F	P	W	I	V	H	4
8	NP_477320.1	Origin recognition complex subunit 4	F	P	-	I	V	H	2
9	NP_477132.1	Origin recognition complex subunit 5	F	P	-	I	V	H	4
10	NP_477319.1	Origin recognition complex subunit 6	-	P	-	I	V	H	5
11	NP_651288.1	polybromo	-	-	W	I	V	H	1
12	NP_476841.1	Recombination repair protein 1	-	P	W	I	V	H	6
13	NP_524274.1	Replication Protein A 70	F	P	-	I	V	H	1
14	NP_523915.1	Replication-factor-C 40 kDa subunit	F	P	W	I	V	H	1
15	NP_525111.1	Ribonucleoside diphosphate reductase small subunit	F	P	W	I	-	H	1
		VI. Cell cycle/Cell proliferation							
1	NP_732309.1	14-3-3epsilon	F	P	-	I	V	H	3
2	NP_476885.2	14-3-3zeta	-	-	W	I	V	H	1
3	NP_001027247.1	CENP-ana	-	-	-	-	-	-	1
4	NP_525053.1	Kinesin-like protein at 3A	-	P	W	I	V	H	2
5	NP_572687.1	Klp10A	-	-	W	I	V	H	1
6	NP_647991.1	mad2	F	P	W	I	V	H	1
7	NP_732105.1	mini spindles	-	P	W	I	V	H	2
8	NP_476683.1	plutonium	-	-	-	-	-	-	1
9	NP_524179.2	polo	F	-	W	I	V	H	4
10	NP_001036712.1	telomere fusion	-	-	-	-	-	-	4
		VII. Chromatin remodelling							
1	NP_536734.2	ATP-dependent chromatin assembly factor large subunit	-	-	W	I	V	H	21
2	NP_611209.1	Brahma associated protein 55 kDa	-	P	-	I	V	H	1
3	NP_649111.1	Chd3	-	-	-	-	-	-	1
4	NP_477128.1	Nucleosome assembly protein 1	F	P	W	I	V	H	1
5	NP_524156.1	reptin	F	P	W	I	V	H	1
		VIII. Transcription							
1	NP_524850.2	lodestar	-	-	W	I	-	H	2
2	NP_477419.1	RNA polymerase II 33 kDa subunit	F	P	W	I	V	H	2
		IX. Translation							
1	NP_001015316.1	CG17514	F	P	W	I	V	H	3
2	NP_523493.2	cup	-	-	-	-	-	-	1
3	NP_733280.1	Ef1gamma	F	P	W	I	V	H	4
4	NP_726411.1	eIF-5A	-	-	W	I	V	H	1
5	NP_723137.1	Eukaryotic initiation factor 4a	-	P	-	I	-	H	4
6	NP_524640.3	Eukaryotic-initiation-factor-4 G	-	-	-	I	V	H	2
7	NP_996433.1	yolkless	-	-	-	I	-	-	1
8	NP_788665.1	B52	-	-	W	I	V	H	1
		X. RNA binding/mRNA processing							
1	NP_648062.2	CG10077	F	P	W	I	V	H	1
2	NP_649645.1	CG1249	F	P	W	I	V	H	1
3	NP_572337.1	CG3198	-	P	W	I	V	H	1
4	NP_651066.2	CG6937	-	P	-	I	V	H	1
5	NP_728652.1	CG7971	-	-	-	-	-	-	3
6	NP_524296.2	hyperplastic discs	-	-	-	I	V	H	2
7	NP_523533.2	maternal expression at 31B	F	P	W	I	V	H	1
8	NP_727962.1	no on or off transient A	-	-	-	I	V	H	1

Table 2. (continued)

S.No.	Accession Number	Protein Name	Homologues*						Total peptide hits
		XI. Transport							
1	NP_648040.1	CG10226	-	P	-	-	V	-	2
2	NP_609967.1	CG10237	-	-	-	-	-	-	1
3	NP_648579.3	CG10657	-	-	-	-	-	-	1
4	NP_523524.2	GDP dissociation inhibitor	F	P	W	I	V	H	2
		XII. Signal Transduction							
1	NP_524631.1	ADP ribosylation factor 102 F	-	-	W	I	V	H	3
2	NP_001097184.1	rho-type guanine exchange factor	-	-	-	-	-	-	3
3	NP_001097440.1	Unc-89	-	-	-	-	-	-	3
		XIII. Mitochondrial							
1	NP_609285.1	Aldehyde dehydrogenase	F	P	W	I	V	H	2
2	NP_726631.1	ATP synthase-beta	F	P	W	I	V	H	2
3	NP_726243.1	bellwether	F	P	W	I	V	H	2
		XIV. Ribosomal							
1	NP_648514.1	Ribosomal protein L10Ab	F	P	W	I	V	H	4
2	NP_728839.1	Ribosomal protein L28	-	-	-	I	V	H	1
3	NP_725114.1	Ribosomal protein S11	F	P	W	I	V	H	2
4	NP_724109.1	Ribosomal protein S26	F	P	W	I	V	H	6
		XV. Others							
1	NP_611483.1	CG13422	-	-	-	-	-	-	1
2	NP_650718.1	CG14309	-	-	-	-	-	-	1
3	NP_610266.1	CG1845	-	-	-	I	V	H	1
4	NP_724671.1	CG30364	-	-	-	-	-	-	2
5	NP_610234.3	CG3420	-	-	-	-	-	-	1
6	NP_650580.1	CG3678	-	P	W	I	V	H	1
7	NP_569839.2	CG3777	-	-	-	-	-	-	1
8	NP_001097950.1	CG4951	-	-	-	-	-	-	1
9	NP_651191.1	CG5857	-	-	-	I	V	H	4
10	NP_650777.1	CG6013	-	P	W	I	V	H	2
11	NP_650693.1	CG7175	-	-	-	-	-	-	2
12	NP_650716.2	CG7671	-	-	-	I	V	H	2
13	NP_611422.2	CG7744	-	-	-	-	V	H	1
14	NP_610810.1	CG8771	-	P	-	I	V	H	1
15	NP_524772.2	dappled	-	-	W	I	V	H	2
16	NP_650993.1	female-specific independent of transformer	-	-	-	-	-	-	1
17	NP_523783.1	GTP-binding-protein	-	P	W	I	V	H	1
18	NP_608690.2	insensitive	-	-	-	-	-	-	1
19	NP_733112.1	jing interacting gene regulatory 1	-	-	-	-	-	-	1
20	NP_610932.1	transport and Golgi organization 7	F	P	W	I	V	H	3
21	NP_476908.1	vacuolar H[+]-ATPase 55kD B subunit	F	P	W	I	V	H	1
22	NP_570008.1	CG8310	-	-	-	-	-	-	4

*F, fungi, including *S. cerevisiae, S. pombe*; P, plants, including *O. sativa and A. thaliana*; W, worms, including *C.elegans*; I, insects, including *A. gambiae*; V, vertebrates including *D. rerio*; H, *H. sapiens*.

consisting of proteins present in both the preparations, has about 50% of single peptide hits, while more dynamic constituents, proteins present in only one preparation, has more than 70% single peptide hits. This may support the idea that dynamic component of the NuMat may be of regulatory nature and, therefore, less abundant, while more

Table 3. Proteins unique to 14–16 h NuMat proteome

S.No.	Accession number	Protein name	Homologues*						Total peptide hits
		I. Nuclear membrane/Nuclear Pores							
1	NP_523588.2	CAS/CSE1 segregation protein	F	P	-	I	V	H	2
2	NP_730182.1	CG32165	F	P	-	I	V	H	1
3	NP_476617.1	Laminin A	-	-	W	I	V	H	4
4	NP_610240.2	klaroid	-	-	-	-	-	-	2
		II. Structural proteins							
1	NP_523800.1	Actin 57B	-	P	W	I	V	H	3
2	NP_524367.1	Actin 88 F	-	P	-	I	-	H	2
3	NP_477484.2	alpha actinin	-	-	W	I	V	H	1
4	NP_476739.1	alpha Spectrin	-	-	W	I	V	H	2
5	NP_524009.2	alpha-Tubulin at 67 C	-	-	-	-	-	-	1
6	NP_523388.1	beta Spectrin	-	-	W	I	V	H	10
7	NP_523842.2	beta-Tubulin at 60D	-	-	-	I	-	H	3
8	NP_477016.1	chickadee	F	P	W	I	-	-	1
9	NP_728438.2	CG17450	-	-	-	I	V	H	1
10	NP_647875.1	Cuticular protein 64Ad	-	-	-	-	-	-	1
11	NP_001097524.1	Cuticular protein 65Ay	-	-	-	-	-	-	1
12	NP_524608.1	gamma-coatomer protein	F	P	W	I	V	H	4
13	NP_511049.1	Myosin light chain cytoplasmic	-	-	W	I	V	H	2
14	NP_476952.1	Microtubule-associated protein 60	-	-	-	-	-	-	2
15	NP_729405.1	Paramyosin	-	-	-	-	-	-	23
16	NP_511057.1	spaghetti squash	-	-	W	I	V	H	1
17	NP_570087.1	Vap-33-1	-	-	W	I	V	H	1
18	NP_477190.1	viking	-	-	-	-	-	-	2
		III. Heat shock proteins/Chaperones							
1	NP_724492.1	lethal (2) 01289	-	-	-	-	-	-	4
		IV. Enzymes/Metabolic process proteins							
1	NP_524996.1	belphegor	P	-	W	I	V	H	1
2	NP_787961.1	CG33128	-	-	-	-	V	-	2
3	NP_609641.1	CG6523	-	-	-	-	-	-	2
4	NP_729177.1	CG8368	-	-	-	-	-	-	4
5	NP_477271.1	Oxysterol binding protein	F	P	W	I	V	H	1
6	NP_788751.1	Papilin	-	-	-	-	V	H	3
7	NP_647836.2	Sc2	F	P	W	I	V	H	1
8	NP_524006.1	Laminin B2	-	-	W	I	V	H	2
		V. DNA binding/Replication							
1	NP_572711.2	CG1737	-	-	-	-	-	-	2
2	NP_609148.1	CG7154	-	-	W	I	V	H	2
3	NP_523529.1	DNA replication-related element factor	-	-	-	-	-	-	14
4	NP_523984.1	Minichromosome maintenance 7	F	P	W	I	V	H	1
5	NP_611806.3	pita	-	-	-	-	-	-	1
6	NP_609494.1	Replication factor C 38 kDa subunit	F	P	W	I	V	H	2
7	NP_524886.2	without children	-	-	-	-	-	-	1
		VI. Cell cycle/Cell proliferation							
1	NP_523374.2	Chromosome-associated protein	F	P	W	I	V	H	1
2	NP_524614.2	modulo	-	-	W	-	-	-	2
3	NP_523430.1	Septin-1	-	-	-	I	V	H	1
4	NP_650048.1	Translationally controlled tumor protein	F	-	W	I	V	H	3

Table 3. (continued)

S.No.	Accession number	Protein name	Homologues*						Total peptide hits
		VII. Chromatin remodelling							
1	NP_730762.2	Chromator	-	-	-	-	-	-	1
2	NP_523441.1	kismet	-	-	W	I	V	H	4
3	NP_725756.1	lola like	-	-	-	-	-	-	1
4	NP_732619.1	modifier of mdg4	F	P	-	I	V	H	2
5	NP_524557.1	Nucleoplasmin	-	-	-	-	-	-	1
6	NP_725841.1	proliferation disrupter	-	-	-	-	-	-	1
		VIII. Transcription							
1	NP_651216.2	CG6129	-	-	W	I	-	H	1
2	NP_649482.2	CG9775	-	-	-	-	-	-	4
3	NP_476605.1	cropped	-	-	-	-	-	-	4
4	NP_523391.3	enhancer of yellow 1	F	-	-	I	V	H	1
5	NP_524846.1	enhancer of yellow 2	-	-	-	-	-	-	1
6	NP_608334.3	enhancer of yellow 3	-	-	-	-	-	-	1
7	NP_524143.2	ftz transcription factor 1	-	-	-	-	-	-	1
8	NP_726567.1	Host cell factor	-	-	-	-	-	-	2
9	NP_523833.1	ken and barbie	-	-	-	-	-	-	2
10	NP_476593.2	lethal (1) 1Bi	-	-	-	-	-	-	2
11	NP_650686.1	Mediator complex subunit 17	-	-	-	I	V	H	1
12	NP_610879.1	Myb-interacting protein 120	-	-	-	I	V	H	1
13	NP_610339.1	spenito	-	-	W	I	V	H	5
14	NP_729571.1	TBP-associated factor 2	-	P	W	I	V	H	1
		IX. Translation							
1	NP_476874.1	string of pearls	F	P	W	I	V	H	2
2	NP_648627.1	SRm160	-	P	W	I	V	H	1
3	NP_611242.1	eIF3-S8	F	P	W	I	V	H	1
4	NP_609543.2	CG5317	-	-	-	-	-	-	4
		X. RNA binding/mRNA processing							
1	NP_995906.1	bancal	-	-	-	-	-	-	1
2	NP_726938.1	cap binding protein 80	-	P	W	I	V	H	1
3	NP_572424.1	CG10777	-	-	-	-	-	-	1
4	NP_723243.1	CG11266	F	P	W	I	V	H	1
5	NP_611354.2	CG30122	-	-	W	-	V	H	2
6	NP_612028.4	CG32344	F	P	W	I	V	H	5
7	NP_611557.2	CG4266	-	-	-	-	-	-	1
8	NP_609356.2	CG4901	-	-	-	I	-	H	2
9	NP_651245.1	CG5728	-	-	-	-	-	-	5
10	NP_648206.1	CG7185	-	-	W	I	V	H	2
11	NP_647626.1	CG7879	-	-	-	-	-	-	1
12	NP_788922.1	CG8611	F	P	-	I	V	H	2
13	NP_610090.1	CG9253	F	P	W	I	V	H	2
14	NP_523434.2	Helicase	F	P	W	I	V	H	1
15	NP_523429.1	penguin	F	P	W	I	V	H	1
16	NP_524446.3	pitchoune	F	P	W	I	V	H	4
17	NP_524858.1	Protein on ecdysone puffs	-	-	-	-	-	-	2

Table 3. (continued)

S.No.	Accession number	Protein name	Homologues*						Total peptide hits
		XI. Transport							
1	NP_610761.2	gartenzwerg	-	P	W	I	V	H	1
2	NP_727448.1	stress-sensitive B	-	-	-	I	V	H	3
		XII. Signal Transduction							
1	NP_511140.1	Casein kinase Ialpha	-	P	-	I	V	H	1
2	NP_647713.2	Dromyosuppressin receptor 1	-	-	-	-	-	-	1
3	NP_476890.1	porcupine	-	-	W	I	V	-	3
4	NP_476857.1	Roughened	-	-	W	I	V	H	1
		XIII. Mitochondrial							
1	NP_572320.1	Mitochondrial assembly regulatory factor	-	-	W	I	V	H	2
2	NP_524358.2	Oligomycin sensitivity-conferring protein	F	P	W	I	V	H	1
		XIV. Ribosomal							
1	NP_609305.1	CG4364	F	P	W	I	V	H	1
2	NP_573099.1	CG8939	F	P	W	I	V	H	6
3	NP_610455.1	Mystery 45A	F	P	W	I	V	H	2
4	NP_611232.1	Ngp	F	P	W	I	V	H	3
5	NP_651976.1	Nnp-1	F	-	-	I	-	H	20
6	NP_730773.3	Qm	-	P	W	I	V	H	5
7	NP_477054.1	Ribosomal protein L11	F	P	W	I	V	H	1
8	NP_477134.1	Ribosomal protein L22	F	P	W	I	V	H	3
9	NP_649070.1	Ribosomal protein L26	F	P	W	I	V	H	2
10	NP_651417.1	Ribosomal protein L27	F	P	W	I	V	H	3
11	NP_524316.1	Ribosomal protein L3	F	P	W	I	V	H	5
12	NP_724149.1	Ribosomal protein L30	F	P	W	I	V	H	2
13	NP_572243.1	Ribosomal protein L35	F	P	W	I	V	H	2
14	NP_001036390.1	Ribosomal protein L5	F	P	W	I	V	H	2
15	NP_651876.1	Ribosomal protein L6	F	P	W	I	V	H	5
16	NP_524726.1	Ribosomal protein L8	F	P	W	I	V	H	4
17	NP_611136.1	Ribosomal protein S15	F	P	W	I	V	H	1
18	NP_727692.1	Ribosomal protein S15Aa	F	P	W	I	V	H	2
19	NP_651359.1	Ribosomal protein S27	F	P	W	I	V	H	2
20	NP_651740.1	Ribosomal protein S8	F	P	W	-	V	H	2
21	NP_524004.2	Ribosomal protein S9	F	P	W	I	V	H	2
		XV. Others							
1	NP_610976.1	CG10139	-	-	-	-	-	-	2
2	NP_649710.1	CG1234	F	P	W	I	V	H	8
3	NP_610347.3	CG12769	-	-	-	-	-	-	2
4	NP_001097279.1	CG13185	F	P	-	I	-	H	3
5	NP_569854.1	CG13362	-	-	-	-	-	-	1
6	NP_650270.1	CG15887	-	-	-	-	-	-	1
7	NP_569860.2	CG16989	-	-	-	I	V	H	2
8	NP_570069.1	CG2875	F	P	-	I	V	H	1
9	NP_572612.2	CG2962	-	-	-	-	-	-	2
10	NP_729720.1	CG32086	-	-	-	-	-	-	1
11	NP_572332.1	CG3226	-	P	-	I	V	H	1
12	NP_001033972.1	CG33988	-	-	-	-	-	-	1
13	NP_609485.2	CG4705	F	-	-	I	V	H	2
14	NP_648889.2	CG4877	-	-	-	-	-	-	2

Table 3. (continued)

S.No.	Accession number	Protein name	Homologues*						Total peptide hits
15	NP_651484.1	CG5468	-	-	-	-	-	-	1
16	NP_651489.1	CG5471	-	-	-	-	-	-	1
17	NP_650124.1	CG6962	-	-	-	I	V	H	2
18	NP_572529.2	CG7065	-	-	-	-	-	-	1
19	NP_651760.1	CG7911	-	-	-	-	-	-	1
20	NP_648406.1	CG8003	-	-	W	I	V	H	2
21	NP_611084.2	CG8414	-	-	-	-	-	-	2
22	NP_649875.1	CG8436	-	-	-	-	-	-	2
23	NP_610484.1	CG8801	F	P	W	I	V	H	3
24	NP_610095.1	CG9246	F	P	W	I	V	H	1
25	NP_649874.1	CG9740	-	-	-	-	-	-	1
26	NP_724341.1	His1	-	P	-	I	-	H	2
27	NP_524519.1	Histone H2A variant	F	-	W	I	V	H	4
28	NP_477379.1	Intronic Protein 259	F	P	W	I	V	H	3
29	NP_649504.1	labial associated factor	-	-	-	-	-	-	1
30	NP_652615.1	meso18E	-	-	-	-	-	-	1
31	NP_649639.1	Osiris 18	-	-	-	-	-	-	1
32	NP_649625.2	Osiris 6	-	-	-	-	-	-	2
33	NP_647975.2	spook	-	-	-	-	-	-	1
34	NP_477033.1	Tiggrin	-	-	-	-	-	-	4
35	NP_650205.1	toys are us	-	P	-	I	V	H	1
36	NP_609589.1	CG5787	-	-	-	-	-	-	2

*F, fungi, including *S. cerevisiae, S. pombe*; P, plants, including *O. sativa and A. thaliana*; W, worms, including *C.elegans*; I, insects, including *A. gambiae*; V, vertebrates including *D. rerio*; H, *H. sapiens*.

abundant core proteome may be contributing to the structural features of the NuMat. For all the further analysis, a list inclusive of proteins identified by both single and two or more than two unique peptides was used. A representative annotated MS/MS spectrum is shown in figure 2.

In order to identify the differential NuMat proteins between the two stages of development, we compared these NuMat proteomes and found that there were 120 proteins unique to the early stage NuMat while 148 proteins were unique to the late stage NuMat. Also, there were 147 proteins common to both the stages. A Venn diagram representing the number of proteins identified in each stage is shown in figure 3. While this differential study indicates that a certain fraction of NuMat proteome is common between different stages and could be interpreted as the minimal or the core NuMat constituent, there is still a significant fraction that is not static and changes in the context of developmental stage.

3.2 *Dynamics of functional classes of NuMat proteome during development*

We categorized NuMat proteins unique to each stage as well as the ones common between the two stages using gene ontology (GO) classification based on molecular functions, biological processes and cellular component (Consortium 2010). The percentage of unique proteins in different functional groups for the particular developmental stage was calculated as mentioned in 'statistical analysis' section. As reported earlier, these functional groups reflected the rich variety of proteins and hence the functions associated with the NuMat (figure 4).

In order to study the complexity and dynamics of NuMat proteome during development, we compared the various functional groups in the two stages in terms of absolute numbers of proteins in each group (tables 2, 3 and 4) instead of percentage abundance as the total number of proteins in both the stages was different. The complexity shift and dynamics of each functional group was calculated as mentioned in 'statistical analysis' section.

According to this analysis, we found that there were six functional groups that showed a trend of increasing complexity from an early stage NuMat to late stage NuMat. These groups included structural proteins, transcription-associated proteins, translation-associated proteins, chromatin-remodelling proteins, signal transduction proteins and ribosomal proteins. Among these groups, transcription-associated proteins showed the maximum dynamics of 89% and the ribosomal proteins showed the minimum dynamics of 42%. The structural proteins, translation-associated proteins, chromatin-remodelling proteins

Table 4. Proteins common between 0–2 h and 14–16 h NuMat proteome

S. No.	Accession number	Protein name	Homologues*						Total peptide hits
		I. Nuclear membrane/Nuclear Pore							
1	NP_477287.1	Nup154	F	P	W	I	V	H	3
2	NP_609446.1	Nup170	-	-	-	I	V	H	8
3	NP_651187.2	Nup98	-	P	-	I	-	H	4
4	NP_572929.1	CG11092	-	-	-	-	-	-	2
5	NP_524226.1	Karyopherin beta 3	F	P	W	I	V	H	1
6	NP_523742.2	Lamin C	-	-	-	-	V	H	3
7	NP_476616.1	Lamin	-	-	W	I	V	H	125
8	NP_726114.1	Lamin B receptor	-	-	-	I	-	-	8
9	NP_477067.2	Megator	F	P	-	I	V	H	4
		II. Structural proteins							
1	NP_511052.1	Actin 5 C	-	P	W	I	V	H	3
2	NP_524210.1	Actin 79B	-	P	-	I	V	H	1
3	NP_477091.1	Actin 87E	-	P	-	I	V	H	1
4	NP_476772.1	alpha-Tubulin at 84B	F	-	W	I	V	H	21
5	NP_523795.2	beta-Tubulin at 56D	F	P	W	I	V	H	12
6	NP_524290.2	beta-Tubulin at 85D	-	P	W	-	V	H	5
7	NP_651606.1	beta-Tubulin at 97EF	-	-	-	I	-	-	22
8	NP_477042.1	Clathrin heavy chain	F	P	W	I	V	H	2
9	NP_723046.1	Collagen type IV	-	-	W	I	V	H	3
10	NP_723999.1	Myosin heavy chain	-	-	W	I	V	H	7
11	NP_001014552.1	zipper	F	-	W	I	V	H	3
		III. Heat shock proteins/Chaperones							
1	NP_650572.2	Cctgamma	F	P	W	I	V	H	5
2	NP_609579.1	CG5525	F	P	W	I	V	H	7
3	NP_572524.1	CG7033	F	P	W	I	V	H	7
4	NP_610418.1	CG8258	F	P	W	I	V	H	2
5	NP_649835.1	CG8351	F	P	W	I	V	H	6
6	NP_651601.1	Glycoprotein 93	-	P	W	I	V	H	2
7	NP_523997.1	Heat shock protein 26	-	-	-	-	-	-	1
8	NP_524000.1	Heat shock protein 27	-	-	-	-	-	-	5
9	NP_511115.2	Heat shock protein 60	F	P	W	I	V	H	2
10	NP_523899.1	Heat shock protein 83 CG1242-PA	-	P	-	I	V	H	4
11	NP_524063.1	Heat shock protein cognate 1	F	P	W	-	V	H	3
12	NP_511132.2	Heat shock protein cognate 3	F	P	W	I	V	H	8
13	NP_788679.1	Heat shock protein cognate 4	F	P	-	I	V	H	41
14	NP_573066.1	lethal (1) G0022	F	P	W	I	V	H	1
15	NP_523707.1	T-complex Chaperonin 5	F	P	W	I	V	H	5
16	NP_732748.1	Tcp1-like	F	P	W	I	V	H	1
		IV. Enzymes/Metabolic process proteins							
1	NP_001027266.1	Alcohol dehydrogenase	-	-	-	-	-	-	1
2	NP_523366.2	Cyclophilin 1	F	P	-	I	V	H	4
3	NP_722721.1	Enolase	F	P	W	I	V	H	2
4	NP_524364.2	FK506-binding protein 1	F	P	-	I	-	-	10
5	NP_001014610.1	GST-containing FLYWCH zinc-finger protein	-	-	-	I	-	-	1
6	NP_572503.1	lethal (1) G0020	F	P	W	I	V	H	2
7	NP_001027227.1	Protein phosphatase 2A at 29B	F	P	W	I		H	3

Table 4. (continued)

S. No.	Accession number	Protein name	Homologues*						Total peptide hits
8	NP_524448.3	Pyruvate kinase	F	P	W	I	V	H	4
9	NP_572463.1	stardust	-	-	-	I	-	-	1
10	NP_727689.1	thioredoxin peroxidase 1	F	P	W	I	V	H	1
11	NP_727078.1	Ubiquitin-5E	-	P	-	I	V	-	1
		V. DNA binding/Replication							
1	NP_001014605.1	Germ line transcription factor 1	F	P	W	I	V	H	2
2	NP_609399.1	RfC3	F	P	W	I	V	H	3
3	NP_001027038.1	terribly reduced optic lobes	-	-	-	I	-	-	4
4	NP_511161.2	Topoisomerase 1	F	P	W	I	V	H	1
5	NP_476760.1	Topoisomerase 2	-	-	-	I	V	H	2
6	NP_649297.1	Z4	-	-	-	I	-	-	1
		VI. Cell cycle/Cell proliferation							
1	NP_001036452.1	Nipped-B	-	P	-	I	V	H	4
2	NP_651211.2	SMC1	F	P	W	I	V	H	2
		VII. Chromatin remodelling							
1	NP_725418.1	Boundary element-associated factor of 32kD	-	-	-	-	-	-	1
2	NP_536746.1	brahma	F	P	W	I	V	H	7
3	NP_728507.1	Enhancer of bithorax	-	-	-	I	-	-	1
4	NP_731927.1	Histone H4 replacement	-	P	W	I	V	-	12
5	NP_523719.1	Imitation SWI	F	-	W	I	V	H	16
6	NP_001014591.1	Mi-2	-	P	W	I	V	H	2
7	NP_524373.1	moira	F	-	W	I	V	H	1
8	NP_001097192.1	Nipped-A	F	P	W	I	V	H	3
9	NP_523849.3	Nucleosome remodeling factor - 38kD	F	-	W	I	V	-	1
10	NP_476755.1	Suppressor of variegation 205	-	-	-	I	V	H	1
11	NP_724749.1	Suppressor of variegation 2-10	-	-	-	I	-	H	1
12	NP_524342.3	Suppressor of variegation 3-7	-	-	-	-	-	-	9
		VIII. Transcription							
1	NP_609389.1	CG5366	-	P	W	I	V	H	1
2	NP_511124.1	RNA polymerase II 215kD subunit	F	P	W	I	V	H	1
		IX. Translation							
1	NP_651605.1	CG4849	F	P	W	-	V	H	4
2	NP_524611.1	Elongation factor 1alpha100E	-	-	-	I	V	H	8
3	NP_477375.1	Elongation factor 1alpha48D	F	-	W	I	V	H	7
4	NP_525105.2	Elongation factor 2b	-	P	W	I	V	-	2
5	NP_726745.2	stubarista	-	P	W	I	V	H	6
		X. RNA binding/mRNA processing							
1	NP_536783.1	belle	F	P	-	I	-	H	4
2	NP_648818.3	CG5931	F	P	W	I	V	H	8
3	NP_524714.1	hoi-polloi	F	P	W	I	V	H	7
4	NP_729237.1	lark	-	-	-	-	-	-	2
5	NP_610735.1	prp8	F	P	W	I	V	H	6
6	NP_524243.2	Rm62	-	P	-	I	V	H	5
7	NP_524774.1	small nuclear ribonucleoprotein at 69D	F	P	W	I	V	H	2
		XI. Transport							
1	NP_001027180.1	CG6783	-	-	-	-	V	H	1

Table 4. (continued)

S. No.	Accession number	Protein name	Homologues*						Total peptide hits
		XII. Signal Transduction							
1	NP_524780.1	moleskin	F	P	-	I	V	H	3
2	NP_732610.1	Rab-protein 1	F	P	W	I	V	H	6
3	NP_727499.1	ran	F	P	W	I	V	H	4
4	NP_524634.2	Retinoid- and fatty-acid binding protein	-	-	-	I	-	-	2
		XIII. Mitochondrial							
1	NP_648905.1	CG4169	-	P	-	I	V	H	1
2	NP_573388.1	lethal (1) G0156	F	P	W	I	V	H	3
		XIV. Ribosomal							
1	NP_523817.1	Fibrillarin	F	P	W	I	V	H	8
2	NP_477412.1	nop5	F	P	W	I	V	H	9
3	NP_651040.3	Nop56	F	P	W	I	V	H	8
4	NP_524819.1	Ribosomal protein L12	-	-	W	I	V	-	1
5	NP_523530.1	Ribosomal protein L13	F	P	W	I	V	H	1
6	NP_649560.1	Ribosomal protein L13A	F	P	W	-	V	H	2
7	NP_523975.1	Ribosomal protein L14	F	P	W	I	V	H	3
8	NP_727119.1	Ribosomal protein L17	F	P	W	I	V	H	4
9	NP_648091.1	Ribosomal protein L18	F	P	W	I	V	H	3
10	NP_523774.1	Ribosomal protein L18A	-	-	-	I	-	-	6
11	NP_476631.1	Ribosomal protein L19	F	P	W	I	V	H	1
12	NP_610144.1	Ribosomal protein L21	-	-	-	I	-	-	1
13	NP_523813.1	Ribosomal protein L23	F	P	W	I	V	H	4
14	NP_523886.1	Ribosomal protein L23A	-	P	W	I	V	H	3
15	NP_609649.1	Ribosomal protein L24	F	P	W	I	V	H	4
16	NP_476963.1	Ribosomal protein L27A	F	P	W	I	V	H	3
17	NP_724805.1	Ribosomal protein L31	F	P	-	I	V	-	1
18	NP_001036440.1	Ribosomal protein L38	-	P	W	I	V	H	2
19	NP_524538.2	Ribosomal protein L4	-	P	W	I	V	H	2
20	NP_523531.1	Ribosomal protein L7	F	P	-	I	V	-	3
21	NP_727094.1	Ribosomal protein L7A	-	P	-	I	-	-	1
22	NP_477161.1	Ribosomal protein L9	-	-	-	I	-	-	1
23	NP_728273.1	Ribosomal protein S10b	-	-	-	-	-	-	3
24	NP_524884.1	Ribosomal protein S14a	F	P	W	I	V	H	2
25	NP_611685.1	Ribosomal protein S16	F	P	W	I	V	H	3
26	NP_524002.1	Ribosomal protein S17	-	-	-	I	V	H	3
27	NP_725943.1	Ribosomal protein S18	F	P	W	I	V	H	2
28	NP_524421.1	Ribosomal protein S20	F	-	W	-	-	H	2
29	NP_610939.2	Ribosomal protein S23	F	P	W	-	V	H	4
30	NP_611693.1	Ribosomal protein S24	F	P	W	I	V	-	9
31	NP_476632.1	Ribosomal protein S3	F	P	W	I	V	H	3
32	NP_524618.1	Ribosomal protein S3A	F	P	W	I	V	H	3
33	NP_729871.1	Ribosomal protein S4	F	P	W	I	V	H	3
34	NP_511073.1	Ribosomal protein S6	F	P	W	I	V	H	1
35	NP_996312.1	Ribosomal protein S7	F	P	W	I	V	H	1
		XV. Others							
1	NP_649652.3	CG1024	-	-	-	-	-	-	1
2	NP_728308.1	CG11943	-	P	W	I	V	H	2
3	NP_572159.1	CG12692	-	-	-	-	-	-	2

Table 4. (continued)

S. No.	Accession number	Protein name	Homologues*						Total peptide hits
4	NP_608340.2	CG14215	-	-	-	-	-	-	4
5	NP_572557.1	CG16892	-	P	-	I	V	H	3
6	NP_730506.1	CG17122	-	-	-	I	V	H	1
7	NP_001036437.1	CG17665	-	P	W	I	V	H	2
8	NP_728599.1	CG2199	-	-	-	-	-	-	1
9	NP_724502.1	CG3287	-	-	-	I	-	-	2
10	NP_611676.1	CG4554	F	P	W	I	V	H	1
11	NP_609493.1	CG4738	-	-	-	-	V	H	5
12	NP_611300.2	CG5733	-	P	W	I	V	H	5
13	NP_651768.1	CG7946	-	-	-	-	-	-	1
14	NP_648367.1	CG8108	-	-	-	I	-	-	1
15	NP_650015.3	fau	-	-	-	I	-	-	1
16	NP_610184.2	gp210	-	-	W	I	V	H	7
17	NP_724343.1	His2A	-	P	-	I	-	H	1
18	NP_001027283.1	His2B	-	-	-	I	V	H	12
19	NP_649899.1	Hrp59	-	-	-	I	V	H	1
20	NP_609079.2	l(2)k09022	-	P	W	I	V	H	9
21	NP_648238.1	rhea	-	-	W	I	V	H	1
22	NP_511103.1	Yolk protein 1	-	-	-	-	-	-	828
23	NP_511102.3	Yolk protein 2	-	-	-	-	-	-	16
24	NP_511148.2	Yolk protein 3	-	-	-	-	-	-	524

*F, fungi, including *S. cerevisiae, S. pombe*; P, plants, including *O. sativa and A. thaliana*; W, worms, including *C.elegans*; I, insects, including *A. gambiae*; V, vertebrates including *D. rerio*; H, *H. sapiens*.

and signal transduction proteins showed the dynamics of 68%, 78%, 48% and 64%, respectively. Although chromatin-remodelling proteins and signal transduction proteins show increasing complexity trend, the complexity shift values are too small to be accounted as significant.

We also found many functional groups that reflected a significant decrease in the complexity between the early and late stage NuMat. These groups included nuclear membrane/nuclear pore proteins, heat shock proteins/chaperones, enzymes/metabolism related proteins, DNA binding/replication proteins, cell cycle/cell proliferation proteins, RNA binding/mRNA processing proteins, transport-associated proteins and mitochondrial proteins. Among these functional groups, cell cycle/cell proliferation proteins showed maximum dynamics of 88%, while heat shock proteins/chaperones showed minimum dynamics of 20%. The rest of the functional groups including nuclear membrane/nuclear pore proteins, enzymes/metabolism-related proteins, DNA binding/replication proteins, RNA binding/mRNA processing proteins, transport-associated proteins and mitochondrial proteins exhibited the dynamics of 55%, 74%, 79%, 71%, 86% and 71%, respectively.

There were many proteins in the NuMat proteome for which we could not assign any function based on GO classification and hence categorized them as 'others'. Although, we have not taken into account the dynamics of this group of proteins for interpreting NuMat dynamics during embryonic development, this does not rule out that these proteins are functionally irrelevant. As seen from figure 4, the proteins in the 'others' group reflect the most complex category of NuMat constituents. Functional characterization of these proteins needs to be done to understand them further in the context of nuclear architecture and embryonic development. Table 1 represents the summary of the complexity and dynamics of the various functional groups of NuMat proteome during development.

We also searched for the conservation of NuMat proteome across the species using NCBI Homologene Database (Sayers *et al*. 2010) and found that many NuMat constituents are conserved across the species (tables 2, 3 and 4). This analysis gives a new angle of study and puts nuclear matrix in the light of evolutionary development.

The classification of proteins in this study is based on the information available in the GO database, which uses known cellular components, biological processes and the molecular functions as the parameters for classification. We find that about one-third of the NuMat proteome unique to early or late embryos and about half of the proteome common to both are composed of proteins known to have

	AA	A	B	Y	
1	L	86.10	114.09	-	20
2	N	200.14	228.13	2036.04	19
3	I	313.22	341.22	1922.00	18
4	T	414.27	442.27	1808.91	17
5	P	511.32	539.32	1707.87	16
6	A	582.36	610.36	1610.81	15
7	T	683.41	711.40	1539.78	14
8	N	797.45	825.45	1438.73	13
9	T	898.50	926.49	1324.69	12
10	A	969.54	997.53	1223.64	11
11	T	1070.58	1098.58	1152.60	10
12	V	1169.65	1197.65	1051.55	9
13	Q	1297.71	1325.71	952.48	8
14	S	1384.74	1412.74	824.43	7
15	F	1531.81	1559.81	737.39	6
16	S	1618.84	1646.84	590.33	5
17	Q	1746.90	1774.90	503.29	4
18	S	1833.93	1861.93	375.24	3
19	L	1947.02	1975.01	288.20	2
20	R	-	-	175.12	1

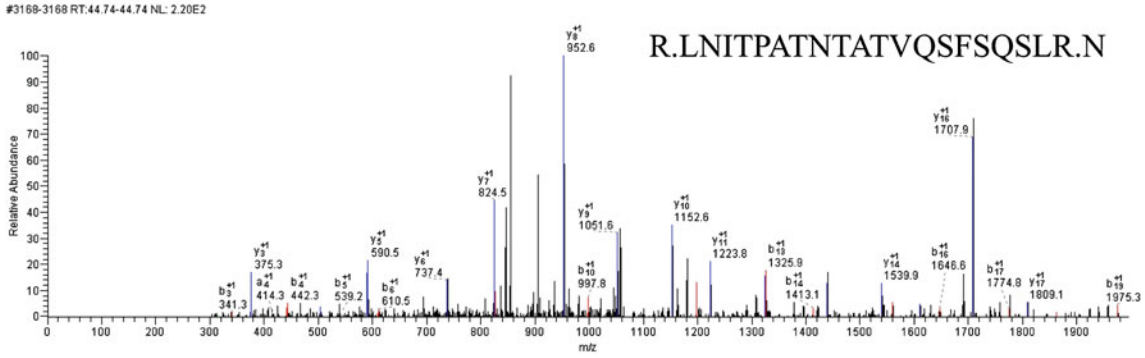

Figure 2. Annotated MS/MS spectrum of Lamin a representative**.** The fragment assignment table and the peptide are shown as an inset in the spectrum. The x-axis and y-axis in the spectrum represent the *m*/*z* and the relative abundance, respectively. The matching A, B and Y ions in the spectrum are highlighted in green, red and blue, respectively, in the fragment assignment table.

cytoplasmic functions. For the nuclei prepared for NuMat proteome analysis, a maximum of 5% cytoplasmic contamination is possible, indicating that NuMat proteome does consist of proteins that have cytoplasmic functions too. It has, indeed, been shown earlier that some cytoplasmic proteins do have nuclear functions (Dion *et al*. 2010; Tariq *et al*. 2009). The role of large proportion of cytoplasmic proteins identified as NuMat component in our analysis, however, remains to be explored.

4. Discussion

In recent years, functional interrelationships between nuclear architecture and gene expression have become increasingly evident (Misteli 2001). Various components of nuclear architecture have provided integrated cues to the mechanistic and functional activities inside the daunting complexity of nuclei. One of these components is the non-chromatin, anastomosing fibrillar meshwork, made of RNA, DNA and proteins, known as nuclear matrix (NuMat) (Capco *et al*. 1982). The regulatory functions of NuMat in processes such as replication (Berezney and Coffey 1975), transcription (S'Iakste and S'Iakste, 2001), mRNA processing (Zeitlin *et al*. 1987), transport of transcripts (Lawrence *et al*. 1989), chromatin loop formation (Dworetzky *et al*. 1992), post-translational modifications of proteins and modification of chromatin structure (Davie 1997) have been implicated in various studies. However, the linkage of various levels of nuclear architecture especially NuMat and the dynamic regulation of gene expression in processes such as development and cellular differentiation has begun to unfold only recently (Chen *et al*. 2001). In this study we show that NuMat proteome from two different stages of development in *D. melanogaster* is remarkably different, which indicates a link between the developmental programme and NuMat.

Embryonic development is a complex and highly orchestrated process. Dramatic changes occurring inside a developing embryo are accompanied by cell-type-specific organization of chromatin and its interaction with proteins inside the nuclear space to bring about differential expression of genes in the context of the specific developmental

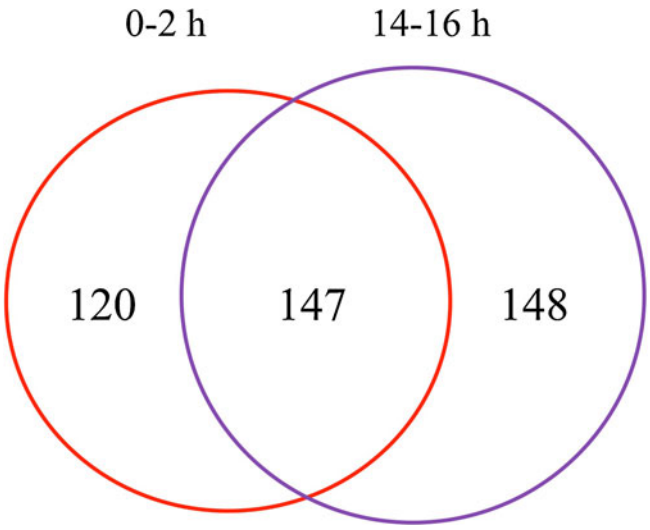

Figure 3. Differential number of proteins in the early- and late-stage NuMat proteome. By LC-MS/MS analysis, the total number of proteins identified in early and late stage NuMat proteome is 267 and 295, respectively. There are 120 proteins unique to the early stage NuMat and 148 proteins unique to the late stage. There are 147 proteins that are common to both the stages.

stage. This dynamics of gene expression puts forward the perspective that the NuMat platform on which these processes are thought to be draped may also alter along with the rest of the nuclear fraction. In fact, this hypothesis is supported by many earlier studies reporting the dynamic nature of nuclear matrix constituents in different cell types. One of the earliest study on the role of nuclear matrix proteins in the development of rabbit granulocytes indicated that nuclear matrix proteins might play a passive structural role in the nuclear changes that accompany granulocyte maturation (Eastment *et al*. 1981). Since then, many studies have shown that nuclear matrix constituents vary in a cell- and tissue-specific manner, which points to the dynamic role of the nuclear matrix in cell differentiation (Bidwell *et al*. 1993; Chen *et al*. 1996; Dworetzky *et al*. 1990; Fey and Penman 1988; Sun *et al*. 1996). In fact, a study on the nucleoskeleton of early bovine embryos suggested that differential expression of some proteins connected with the nucleoskeleton could be used to monitor nuclear reprogramming (Degrouard *et al*. 2004).

The dynamics of *Drosophila* development starting from fertilization to adult fly is a well-studied subject. However, the control of spatiotemporal expression of genes during this process in the context of nuclear architecture still remains to be explored. The comparative analysis of the NuMat proteomes of early and late stage embryos that we report here provides cues to the role of NuMat in embryonic development of *Drosophila.* Our analysis shows that 65% of the whole NuMat proteome is different between the two developmental stages and only 35% remains static, which can be interpreted as the core NuMat proteome.

As seen from table 1, the functional class that show significant increase in the complexity as the embryo develops includes structural proteins, transcription-associated proteins, translation-related proteins and ribosomal proteins. These functional class of proteins fall into the category of 'late complexity'. Within this category, transcription-associated proteins exhibit the 'most dynamic' nature and the ribosomal proteins exhibit the 'least dynamic' nature. The proteins associated with the process of transcription and translation show 'highly dynamic' pattern. In *Drosophila* embryos, transcription starts after the 10th nuclear division (Pritchard and Schubiger 1996). This indicates that a late embryo has more variety of transcription-associated proteins as well as protein machineries required for the translation of the transcripts as compared to the early embryo. It is also known that differentiated cells are structurally and functionally more complex as compared to the non-differentiated cells of the early embryo. These observations explain why NuMat proteome of late stage embryo has higher complexity and dynamics of these particular classes of proteins.

The functional class such as nuclear membrane/nuclear pore, enzymes/metabolism-associated proteins, heat shock proteins/chaperones, DNA binding/replication proteins, cell cycle/cell proliferation proteins, RNA binding/mRNA processing and transport-associated proteins group together as 'early complexity' category. Within this category, cell cycle/cell proliferation proteins exhibit the 'most dynamic' and the heat shock proteins/chaperones display the 'least dynamic' nature. The 'moderately dynamic' class among these functional groups includes nuclear membrane/nuclear pore proteins. The rest of the functional classes in this

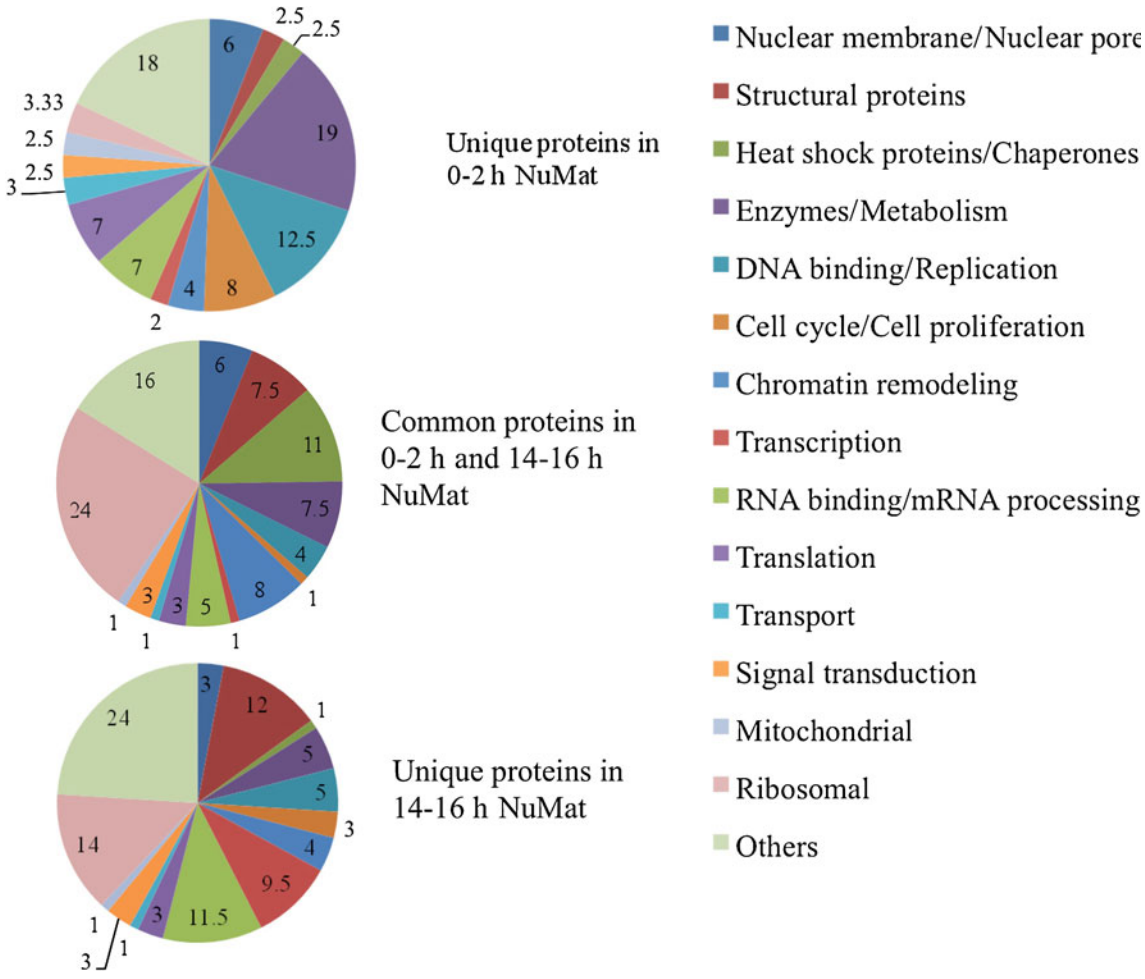

Figure 4. Comparative analysis of the functional classes of the early- and late-stage NuMat. The proteins unique to early and late stage NuMat proteomes as well as the ones common to the two stages are divided into 15 functional categories. The abundance of each category of proteins varies between the two stages.

category such as enzymes/metabolism-associated proteins, DNA binding/replication proteins, RNA binding/mRNA processing and transport-associated proteins exhibit 'highly dynamic' nature. The development of single-celled embryo into a multicellular organism is the result of a series of mitotic divisions that is intricately modulated during the entire course of development from an early embryo to the terminally differentiated cells. This indicates that the kind of proteins associated with cell cycle progression may also differ remarkably between the two developmental stages. In fact, many studies have shown that developmental decisions and fate determination can be controlled by key cell cycle regulators at specific developmental stage (Budirahardja and Gonczy 2009). In the present study also, the NuMat proteomes of two developmental stages reflect this difference in the complexity and dynamics of these classes of proteins and suggest possible coupling of NuMat and cell cycle regulators during the process of embryonic development. Although, mitochondrial proteins are seen as 'early complexity, highly dynamic' class in our analysis, it is not clear whether they constitute bonafide NuMat components or are cytoplasmic contaminants that may be co-purifying with the NuMat fraction.

The functional classes such as chromatin-remodelling proteins and signal transduction proteins displayed a trend towards 'late complexity' but were not considered true members of this category due to lack of significant shifts values. However, both chromatin-remodelling and signal transduction proteins display 'moderately dynamic' nature in context of the NuMat and embryonic development. The process of development requires tremendous changes in the chromatin organization and its association with various proteins in the nuclear space. These changes are mediated with the help of chromatin-remodelling proteins (Ho and

Crabtree 2010). The dynamics of this class of proteins in the NuMat proteome indicates that although there are a few stage-specific chromatin-remodelling proteins that have non-redundant developmental functions, there are many that are common to both the stages and might be involved in most of the basic chromatin activities.

The most complex category of NuMat constituents in early and late stage proteomes is the 'others', the proteins that could not be grouped into any functional class due to lack of specific functional attributes. Since this is a 'highly dynamic' group, it suggests that large number of proteins with a variety of functions contribute to the dynamics of the NuMat proteome. We would also like to emphasize that we are yet to obtain a complete list of NuMat proteins, and therefore, a more extensive list and analysis of more number of developmental stages may lead to a more complex picture. Although it is difficult to achieve, a number of cell type specific NuMat proteome analysis can provide a more complete view of the dynamic nature and functional links of NuMat components with the developmental programme.

In conclusion, our report shows that NuMat proteome of two different stages is remarkably different and shifts in the complexity and the dynamics of the functional classes of NuMat proteome correlates well with the dynamics of the molecular events during development. As indicated above, however, a lot more molecular and biochemical work is required to get a better understanding of this dynamics and to explain the role of NuMat in embryonic development. Our study is the first large-scale analysis that explores the trends in the dynamics of NuMat during development. More importantly, this study identifies a large number of proteins that show remarkable dynamics as NuMat constituents and opens the flood gates of studies that can employ genetics, cell and molecular biology techniques to understand the molecular interactions and mechanisms involved in this complex process. It is now known that certain NuMat constituents are specific for some pathological conditions and are being used as markers for diagnosis of the diseases (Davido and Getzenberg 2000). An analysis of the NuMat proteome in developmental stages also has potential to provide key candidates whose role can be dissected for the diagnosis and therapy of development-related disorders.

Acknowledgements

We thank the CCMB proteomics facility for help in the proteomics experiments. We thank Sreenath for maintenance of fly cages and help in embryo collection. PV acknowledges the Senior Research Fellowship from Council for Scientific and Industrial Research (CSIR), India, and RKM acknowledges grants from Department of Biotechnology and Department of Science and Technology, India.

References

Albrethsen J, Knol JC and Jimenez CR 2009 Unravelling the nuclear matrix proteome. *J. Proteomics* **72** 71–81

Anachkova B, Djeliova V and Russev G 2005 Nuclear matrix support of DNA replication. *J. Cell Biochem.* **96** 951–961

Berezney R and Coffey DS 1975 Nuclear protein matrix: association with newly synthesized DNA. *Science* **189** 291–293

Berezney R and Coffey DS 1977 Nuclear matrix. Isolation and characterization of a framework structure from rat liver nuclei. *J. Cell Biol.* **73** 616–637

Berezney R and Wei X 1998 The new paradigm: integrating genomic function and nuclear architecture. *J. Cell Biochem. Suppl.* **30–31** 238–242

Bidwell JP, Van Wijnen AJ, Fey EG, Dworetzky S, Penman S, Stein JL, Lian JB and Stein GS 1993 Osteocalcin gene promoter-binding factors are tissue-specific nuclear matrix components. *Proc. Natl. Acad. Sci. USA* **90** 3162–3166

Budirahardja Y and Gonczy P 2009 Coupling the cell cycle to development. *Development* **136** 2861–2872

Capco DG, Wan KM and Penman S 1982 The nuclear matrix: three-dimensional architecture and protein composition. *Cell* **29** 847–858

Chen HY, Sun JM, Hendzel MJ, Rattner JB and Davie JR 1996 Changes in the nuclear matrix of chicken erythrocytes that accompany maturation. *Biochem. J.* **320 Pt 1** 257–265

Chen JL, Guo SH and Gao FH 2001 Nuclear matrix in developing rat spermatogenic cells. *Mol. Reprod. Dev.* **59** 314–321

Consortium 2010 The Gene Ontology in 2010: extensions and refinements. *Nucleic Acids Res.* **38** D331–D335

Cremer T, Kupper K, Dietzel S and Fakan S 2004 Higher order chromatin architecture in the cell nucleus: on the way from structure to function. *Biol. Cell* **96** 555–567

Davido T and Getzenberg RH 2000 Nuclear matrix proteins as cancer markers. *J. Cell Biochem. Suppl.* **35** 136–141

Davie JR 1997 Nuclear matrix dynamic histone acetylation and transcriptionally active chromatin. *Mol. Biol. Rep.* **24** 197–207

Degrouard J, Hozak P, Heyman Y and Flechon JE 2004 Nucleoskeleton of early bovine embryos and differentiated somatic cells: an ultrastructural and immunocytochemical comparison. *Histochem. Cell Biol.* **121** 441–451

Dion V, Shimada K and Gasser SM 2010 Actin-related proteins in the nucleus: life beyond chromatin remodelers. *Curr. Opin. Cell Biol.* **22** 383–391

Dworetzky SI, Fey EG, Penman S, Lian JB, Stein JL and Stein GS 1990 Progressive changes in the protein composition of the nuclear matrix during rat osteoblast differentiation. *Proc. Natl. Acad. Sci. USA* **87** 4605–4609

Dworetzky SI, Wright KL, Fey EG, Penman S, Lian JB, Stein JL and Stein GS 1992 Sequence-specific DNA-binding proteins are components of a nuclear matrix-attachment site. *Proc. Natl. Acad. Sci. USA* **89** 4178–4182

Eastment CE, Scott RB, Shelton KR and Haar JL 1981 The role of the nuclear protein matrix during development of rabbit granulocytes. *Blood* **57** 747–757

Eng JK, McCormack AL and Yates III, JR 1994 An approach to correlate tandem mass spectral data of peptides with

amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrometry* **5** 976

Fey EG and Penman S 1988 Nuclear matrix proteins reflect cell type of origin in cultured human cells. *Proc. Natl. Acad. Sci. USA* **85** 121–125

Ho L and Crabtree GR 2010 Chromatin remodelling during development. *Nature (London)* **463** 474–484

Jackson DA 2003 The principles of nuclear structure. *Chromosome Res.* **11** 387–401

Jackson DA and Cook PR 1988 Visualization of a filamentous nucleoskeleton with a 23 nm axial repeat. *EMBO J.* **7** 3667–3677

Kallappagoudar S, Varma P, Pathak RU, Senthilkumar R and Mishra RK 2010 Nuclear matrix proteome analysis of *Drosophila melanogaster*. *Mol. Cell Proteomics* **9** 2005–2018

Lanctot C, Cheutin T, Cremer M, Cavalli G and Cremer T 2007 Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat. Rev. Genet.* **8** 104–115

Lawrence JB, Singer RH and Marselle LM 1989 Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell* **57** 493–502

Misteli T 2001 Protein dynamics: implications for nuclear architecture and gene expression. *Science* **291** 843–847

Pathak RU, Rangaraj N, Kallappagoudar S, Mishra K and Mishra RK 2007 Boundary element-associated factor 32 B connects chromatin domains to the nuclear matrix. *Mol. Cell Biol.* **27** 4796–4806

Pritchard DK and Schubiger G 1996 Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev.* **10** 1131–1142

Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, Dicuccio M, Federhen S**,** *et al*. 2010 Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **38** D5–16

S'Iakste NI and S'Iakste TG 2001 Transcription factors and the nuclear matrix. *Mol. Biol. Mosk.* **35** 739–749

Sun JM, Chen HY, Litchfield DW and Davie JR 1996 Developmental changes in transcription factors associated with the nuclear matrix of chicken erythrocytes. *J. Cell Biochem.* **62** 454–466

Tariq M, Nussbaumer U, Chen Y, Beisel C and Paro R 2009 Trithorax requires Hsp90 for maintenance of active chromatin at sites of gene expression. *Proc. Natl. Acad. Sci. USA* **106** 1157–1162

Zeitlin S, Parent A, Silverstein S and Efstratiadis A 1987 Pre-mRNA splicing and the nuclear matrix. *Mol. Cell Biol.* **7** 111–120

*e*Publication: 08 July 2011