Identification of genes expressed during *Drosophila melanogaster* gastrulation by using subtractive hybridization

Mauricio González-Agüero, Alejandro Zúñiga, Hans Pottstock, Talía del Pozo, Mauricio González, Verónica Cambiazo*

Laboratorio de Bioinformática y Expresión Génica, INTA, Universidad de Chile. Macul 5540, Santiago, Chile

Abstract

A subtractive hybridization approach was used to identify genes that are expressed at the beginning of gastrulation. We used tester DNA complimentary to RNA (cDNA) prepared from stages 6–7 embryos (gastrula) and excess driver cDNA from stages 2–4 embryos (syncytial blastoderm) to generate a gastrula-subtracted cDNA library. A reverse Northern blot procedure used to analyze 105 subtracted clones showed that 65% had a level of expression at least 2.5-fold higher in stages 6–7 versus stages 2–4 embryos. We determined the nucleotide sequence of these clones and identified 49 individual sequences, including 33 previously uncharacterized genes. We verified the level of expression of 24 genes during *Drosophila melanogaster* embryogenesis using a semiquantitative polymerase chain reaction (PCR) approach. As expected, all of the selected clones showed their highest level of expression after stages 2–4 of embryogenesis, including several that displayed peaks of expression during gastrulation. Three genes that were expressed at their highest levels in stages 6–7 were further analyzed by 5′-rapid amplification of cDNA ends (RACE) analysis, Northern blot assays and in situ hybridization. Our results indicate that these genes exhibited temporal and spatially restricted patterns of expression in developing embryos, and moreover, their transcripts were detected in cells that undergo morphological changes during the gastrulation stage. Characterizing the role of these genes will be important to increase our understanding of the molecular mechanisms that regulate cellular activities during *D. melanogaster* gastrulation.

Keywords: Gene expression; Sala; CG13427; CG6234; Embryonic development

1. Introduction

During the early phase of *Drosophila melanogaster* embryonic development, cellular blastoderm is formed, and gastrulation takes place. Throughout the cellular blastoderm stage, the *D. melanogaster* embryo consists of an epithelium of approximately 6000 cells, all of which, with the exception of the germ cells, are uniform in size and shape. Shortly after blastoderm cellularization is completed, several coordinated morphogenetic processes that initiate the gastrulation stage drive an extensive reorganization of the embryonic epithelium (Campos-Ortega and Hartenstein, 1985; Costa et al., 1993). First, mesoderm is internalized through the formation of the ventral furrow. The endoderm then invaginates from each end of the embryo becoming the anterior and posterior midgut. Concomitant with posterior midgut invagination, germ band extension drives cells around the posterior tip of the embryo. As germ band extension occurs, cells at the dorsal midline differentiate

Abbreviations: ds, double-strand; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; SSH, supression subtractive hybridization; SMART, Switching Mechanism At 5' end of RNA Template; cDNA, DNA complimentary to RNA; mRNA, messenger RNA; EST, expression sequence tag; dNTP, deoxyribonucleoside triphosphate; DTT, dithiothreitol; EGTA, ethylene glycol tetraacetic acid; bp, base pair; kb, kilobase; aa, amino acid; BLAST, Basic Local Alignment Search Tool; DIG, digoxigenine; PBS, phosphate buffer saline; CG, Computed Gene; BDGP, Berkeley *Drosophila* Genome Project.

* Corresponding author. Tel.: +56 2 6781514; fax: +56 2 2214030.
E-mail address: vcambiaz@inta.cl (V. Cambiazo).

into the amnioserosa. Several hours later, the germ band will retract back around the posterior tip, repositioning the caudalmost body parts at the posterior end of the embryo. Concurrently with these morphogenetic processes, the expression of zygotic transcripts is activated for the first time at the beginning of cellular blastoderm formation and subsequently at the gastrulation stage when degradation of the maternal transcripts has already started (Wieschaus and Sweeton, 1988).

Our current knowledge of genes that play roles during *D. melanogaster* development is derived from the analyses of embryonic lethal mutants that display defects in larval cuticle morphology. Such genetic strategies have illuminated the mechanisms by which cells are assigned positional values and are fated to contribute to particular tissue types (Wieschaus, 1996). These studies have allowed us to recognize genes that are essential for embryonic patterning, whereas their molecular characterization has revealed that they code for components of signal transduction pathways and their associated transcription factors. However, our understanding of the gene networks that control cellular behavior during gastrulation is incomplete. This points to the limitations of classic genetic techniques since mutants with a partially penetrant phenotype may be ignored and functionally redundant genes are entirely missed (Miklos and Rubin, 1996). Few genes have been found, which encode proteins with structural or regulatory functions during the cellular shape changes of gastrulation (Leptin, 1999). One example is *folded-gastrulation* (*fog*), which encodes a putative secreted protein that is expressed in the invagination primordia of ventral furrow and posterior midgut in a pattern that precisely prefigures the pattern of apical cell constrictions (Costa et al., 1994). *fog* mutations do not affect the assignment of positional values to individual cells, but rather they specifically disrupt coordinate cell shape changes.

An alternative to the genetic approach is the isolation of genes that are expressed at the beginning of the gastrulation stage as it would be expected that genes encoding proteins with structural or regulatory functions would be activated at that time in cells that form the primordium of the germ layers. As a first step towards achieving this, we applied a supression subtractive hybridization (SSH) procedure using tester DNA complimentary to RNA (cDNA) prepared from stages 6–7 embryos (gastrula) and driver cDNA from stages 2–4 embryos (syncytial blastoderm). We generated a cDNA library from gastrula-subtracted transcripts, and sequence analyses were conducted to identify clones of genes that were expressed at a higher level in gastrula than in syncytial blastoderm. A semiquantitative polymerase chain reaction (PCR) screening of 24 clones showed the dynamic expression pattern of these genes during five stages of *D. melanogaster* embryogenesis. In addition, we describe in further detail the transcripts of three zygotic genes that showed the highest increment in their relative abundance during gastrulation. They were characterized by 5′-rapid amplification of cDNA ends (RACE) analysis, Northern blot and in situ hybridization.

2. Materials and methods

2.1. Fly culture and embryo selection

D. melanogaster (Canton-S strain) flies were grown at 22 °C on standard cornmeal, molasses, agar and yeast medium. To obtain staged embryos, females were allowed to lay eggs for 2 h on 2% apple juice agar plates spread with live brewer's yeast. The plates were replaced several times, and finally 1-h embryos were collected, dechorionated and washed with Ringer *Drosophila* solution (182 mM KCl, 46 mM NaCl, 3 mM $CaCl_2$, 10 mM Tris–Hcl, pH 7.2). Embryos were selected at stages 2–3 (syncytial blastoderm) or 5 (cellular blastoderm) and then allowed to continue their development in a humidified chamber at 25 °C. We hand-selected embryos at different developmental stages based on their morphological characters (Campos-Ortega and Hartenstein, 1985) and rapidly froze them in liquid N_2. Embryos were kept at −80 °C for 1–2 weeks.

2.2. RNA extraction and cDNA synthesis

Total RNA was extracted from staged embryos (n=100) using the RNA_{WIZ} reagent (Ambion). Embryos were carefully homogenized in a 1.5 mL eppendorf tube with 1 mL of RNA_{WIZ} reagent using a plastic tissue grinder. To improve RNA yield, the homogenized was passed through Qiashreder columns (Qiagen) by centrifugation at room temperature for 2 min at 10,000 ×g. The quantity and quality of the RNA were assessed by $OD_{260/280}$ and by electrophoresis on a 1.2% formaldehyde-agarose gel. Typical yield was 0.18–0.23 μg RNA/embryo. For the subtractive hybridization procedure, 0.5 μg of total RNA from stages 2–3 and stages 6–7 embryos was used to produce double-strand (ds) cDNA using the Switching Mechanism At 5' end of RNA Template (SMART)™ PCR cDNA Synthesis Kit (Clontech) after 17 cycles of amplification. For semiquantitative PCR assays, 2 μg of total RNA was used to synthesize cDNA with Superscript II reverse transcriptase (Invitrogene).

2.3. Subtractive hybridization

A subtracted cDNA library was made using driver ds-cDNAs from stages 6–7 embryos (gastrula) and tester ds-cDNA from stages 2–3 embryos (syncytial blastoderm) We performed the cDNA subtraction using the PCR-select cDNA subtraction kit (BD Clontech) according to manufacturer's recommendations with modifications. Briefly, we used a driver/tester ratio of 2:1 (volume/volume) in the first hybridization. We carried out 25 cycles of primary PCR and 12 cycles of secondary PCR with the Advantage cDNA polymerase mix (BD Clontech). To evaluate the efficiency

of the cDNA subtraction, we compared the transcript levels of tubulin and actin by RT–PCR in subtracted and unsubtracted cDNA populations of embryos at stages 6–7. Detection of the PCR products of these transcripts required 25 PCR cycles with subtracted cDNA as template, whereas, as expected, only 20 cycles were sufficient to amplify them from control cDNAs. Furthermore, to assess the efficiency of enrichment of differentially expressed genes, the abundance of the transcription factor *twist* was examined. As expected, transcripts of *twist* were enriched in the gastrula-subtracted sample as compared with the unsubtracted sample. The *twist*-specific product was detectable after 20 cycles when subtraction had been carried out but not until 25 cycles in the corresponding unsubtracted sample. As described by the subtraction kit manufacturer (BD Clontech), we estimate at least a 20-fold enrichment in the subtracted cDNA population (data not shown).

Aliquots of 100 μL of secondary PCR reaction from gastrula-subtracted cDNA sequences were purified using the Wizard DNA Clean-up system (Promega) and resuspended in 50 μL of water; 1 μL was inserted into the T/A cloning vector pGEM-T Easy (Promega) following manufacturer's recommendations. Individual transformants carrying cDNA fragments were isolated from white colonies on X-gal/IPTG agar plates and placed in an arrayed 96-well format. Plasmid DNA preparations were made from 105 cDNA clones by alkaline lyses, digested with the *Eco*RI restriction enzyme and analyzed in 1.2% agarose gel electrophoresis. The entire set of clones contained inserts with an average size of 400 base pair (bp).

Sequencing was carried out manually using the dideoxynucleotide chain termination procedure (Ausubel et al., 1997). Sequence homology searches were done against *D. melanogaster* sequence database, using the Basic Local Alignment Search Tool (BLAST)N program at FlyBase (http://flybase.net/blast/) and the BLASTN and BLASTX at NCBI (http://www.ncbi.nlm.nih.gov/BLAST). Domain-based analyses used SMART (http://smart.embl-heidelberg.de) and Interpro (http://www.ebi.ac.uk/interpro/). Signal peptide prediction used SignalP (http://www.cbs.dtu.dk/services/SignalP/). Sequence management was performed with database tools developed in our laboratory.

2.4. cDNA arrays and Northern blot analysis

Plasmid DNA preparations were amplified by PCR using the T7/SP6 forward/reverse primers. The PCR products were purified by salt precipitation, recovering 60% of DNA, and arrayed in 96-well plates. DNA (60 ng) was spotted on a positive charged Nylon membrane (Zeta-Probe, Bio-Rad) using a 96-pin multiblot replicator (V&P Scientific). Each clone plus control cDNAs were spotted in quadruplicate keeping five neighbor spots free of cDNA to determine local background. The filters were washed twice with denaturing solution (0.5 M NaOH, 1.5 M NaCl) and then with neutralizing solution (0.5 M Tris–HCl, pH 7.4, 1.5 M NaCl). After that, DNA was permanently attached to the membrane by UV cross-linking. Hybridization conditions for reverse Northern blot analysis were as follows: prehybridization in 10 ml of buffer 2× prehybridization/hybridization solution (Life Technologies) plus 0.2× sodium phosphate buffer, pH 6.5, 40% formamide and 1% SDS for 1–3 h at 42 °C. Hybridization was performed in the same buffer for 16–18 h at 42 °C. Membranes were washed at room temperature with two 10-min changes of 2× SSC and 2× SSC plus 0.1% SDS, then with two 15-min changes of 1× SSC plus 0.1% SDS at 60 °C and finally with two 15-min changes of 0.1× SSC. Hybridization spots were visualized by exposure to X-ray films at −80 °C for 1 to 3 days. The intensity of the spots was recovered with the Kodak Digital Science 1D software. Probes that corresponded to double-strand SMART cDNAs from stages 2–3 and stages 6–7 embryos were labeled with ^{33}P using the Random Primer Labeling Method (Invitrogene), and unincorporated radioactive nucleotides were removed using the Quiquick Nucleotide removal kit (Quiagen). We performed duplicate hybridization experiments. The average intensity of the four spots for each gene minus the average of local background intensity was normalized using the values obtained for the housekeeping genes (actin and tubulin). Only spots with hybridization intensity two times above the local background measurements were included in the analyses (Tapia et al., 2004).

For conventional Northern blot analysis, total RNA (8 μg) from staged embryos was fractionated in 1% agarose gels containing 6% formaldehyde. Gels were transferred to Nylon membranes following standard procedures (Ausubel et al., 1997). Hybridization, washing and blot exposure conditions were as described above. Probes corresponding to the DNA inserts were labeled with ^{32}P using the Random Primer Labeling Method (Invitrogene).

2.5. Semiquantitative PCR

Aliquots of cDNA (0.2 μg/μl) from embryos at different stages of development: 2–3 (15–90 min), 5 (150–195 min), 6–7 (195–225 min), 8–9 (225–310 min) and 10–12 (310–630 min) were amplified with 1 U of Taq DNA polymerase (Invitrogene) in the presence of specific primers for each selected gene (10 to 50 μM) and tubulin, which was used as an internal control. Primers were designed using the software Primer Premier 3 and synthesized by Invitrogene. Reactions were carried out in a PTC-100 thermocycler (MJ Research). The standard program was comprised of 30 s at 94 °C, 30 s at 56 °C and 2 min at 68 °C for a number of cycles previously determined to assure that amplification was in the exponential range. The PCR products were analyzed in a 1.2% agarose gel; images of the ethidium-bromide-stained agarose gels were acquired with a Kodak DC-290 camera, and quantification of the bands was performed by Kodak 1 D, v. 3.5 program. The intensity of

the amplified products was normalized to the control values of tubulin. For each gene, 100% corresponded to its maximum intensity at a given stage of development.

2.6. 5′-RACE–PCR

5′-RACE analysis was carried out using the SMART cDNA system (Clontech) as described by Gong and Ge (2000), with the following modifications: 1 μg of total RNA of embryos at stages 6–7 was incubated at 70 °C for 2 min with 10 μM oligo (dT) and 12 μM SMART™ II oligonucleotide (5′-AAGCAGTGGTATCAACGCAGAG-TACGCGGG-3′), which became incorporated onto the 5′-end of the reverse-transcribed cDNA. After addition of 200 U of Superscript II reverse transcriptase plus 2 μL of 5× first-strand buffer, 0.01 M dithiothreitol (DTT) and 1 μM deoxyribonucleoside triphosphates (dNTPs), the reaction was carried out for 1 h at 42 °C. The RNA was digested using 1 U of RNase H (Invitrogene). One hundred nanograms of the sscDNA were amplified by primary PCR using 0.48 μM of SMART PCR primer (5′-AAGCAGTGGTAT-CAACGCAGAGT-3′) and 0.40 μM of gene specific primers (CG13427-a: 5′-CTACCAGTAGCGTCGG-3′ or CG6234-a: 5′-GAT TTCCGCCGTAAACCGAT-3′). Secondary PCR was carried out using 40 ng of primary PCR as template and the same SMART PCR primer together with 0.40 μM nested gene-specific primers (CG13427-b: 5′-CAGTAGCGTCGGCCCTGGT-3′ or CG6234-b: 5′-TGCTCAGTGATGGCTTTGTC-3′). The PCR program was comprised of 30 s at 94 °C, 30 s at 62 °C, 2 min at 68 °C for 20 to 28 cycles and a final extension of 68 °C for 5 min. Amplified 5′-RACE fragments were analyzed by agarose gel electrophoresis, and selected bands were purified using the QIAquick Gel Extraction kit (QIAgene). DNA fragments were cloned into pGEM-T easy vectors (Promega) according to manufacturer's recommendations, and both strands were sequenced.

2.7. In situ hybridization of whole-mount embryos

In situ hybridization using digoxigenine (DIG)-labeled DNA probes was carried out as described by Lehmann and Tautz (1994), with the following modifications: methanol–ethylene glycol tetraacetic acid (EGTA) washed embryos were postfixed in a gradual series of methanol–EGTA and 4% paraformaldeyde solution (PP) followed by 20-min incubation in PP and rinsed with phosphate buffer saline (PBS), 0.1% Tween (PBT). The embryos were digested with 2.5 μl of proteinase K (50 μg/ml) in 1 mL of PBS for 4 min. Embryos were incubated with alkaline-phosphatase-conjugated anti-digoxigenin antibody (Boehringer Mannheim) 1:2000 at room temperature for 3 h. The embryos were then washed extensively with PBT, and expression patterns were visualized by incubating them with staining solution containing NBT and BCIP as substrates. The reaction was stopped by washing the samples with 20 mM PBT–EDTA. The stained embryos were dehydrated in a series of ethanol and xylene, mounted in Cytoseal ™ XYL (Richard-Allan Scientific) and photographed on a Zeiss Axiovert 25 microscope.

3. Results and discussion

3.1. Analysis of the subtracted cDNA library

To isolate transcripts that are differentially expressed between gastrulation and syncytial blastoderm stages, we subtracted tester cDNA prepared from stages 6–7 with an excess of driver cDNA from embryos at stages 2–3 using the SSH procedure (Diatchenko et al., 1996). Since the great majority of maternally supplied messenger RNAs (mRNAs) are present at this stage, we expected to deplete these transcripts to identify zygotic mRNAs, including mRNAs with peak of expression at gastrulation.

Subtractive hybridization procedures used to investigate differential expression of genes during *D. melanogaster* development have been previously focused on the identification of genes involved in mesoderm specification and muscle differentiation (Casal and Leptin, 1996; Taylor, 2000). Thus, they have paid attention to genes expressed at postgastrulating stages in the mesoderm and its derivatives (Taylor, 2000). To our knowledge, the approach adopted here has for the first time yielded a cDNA population enriched in transcripts corresponding to genes expressed early during gastrulation.

As a first step in the characterization of our gastrula-subtracted cDNA library, we arrayed 105 clones in duplicate filter sets and screened them with radioactive probes. Replicate double-strand cDNA probes were synthesized from gastrula and syncytial blastoderm RNA preparations and hybridized to the filter arrays made from the subtracted cDNA clones. After background-subtraction and normalization, hybridization signal intensities that consistently differed by more than 2.5-fold between gastrula and syncytial blastoderm stages were considered as corresponding to genes differentially expressed between these two stages. The results of the hybridizations showed that 65% of the clones (n=68) fulfilled this filter criterion and that they were selected for sequencing. After sequence analysis, nonredundant genes (n=49) were grouped according to the function of their encoded proteins (Table 1). Functionally uncharacterized genes, named as novel in Table 1, were classified according to their shared protein domain similarities with known proteins in databases.

3.2. Genes differentially expressed between gastrulation and syncytial blastoderm stages

For each gene, Table 1 displays the name, functional characteristic and Computed Gene (CG) number, which links to FlyBase (http://flybase.net). Among the differentially expressed genes, we found developmentally regu-

Table 1
Genes identified from the gastrula-subtracted cDNA library

CG number	Gene name	Characteristics
Nucleic acid binding proteins		
CG16738	*sloppy paired 1*, *slp1*	Fork head domain
CG3851	*odd skipped*, *odd*	Zinc finger (C2H2)
CG2956	*twist*, *twi*	bHLH domain
CG17334	*lin-28*	Zinc finger (CCHC)
CG4547	Novel	Ataxin domain
CG7357	Novel	Zinc finger (C2H2)
CG17829	Novel	Zinc finger domain, U1-like
Secreted and transmembrane proteins		
CG6736	*insulin-like peptide 4*, *Ilp4*	Insulin/IGF/relaxin family
CG4922	*spalt-adjacent*, *sala*	Secreted protein
CG8827	*Angiotensin converting enzyme*, *Ance*	Metalloprotease
CG13333	Novel	Putative secreted protein
CG13427	Novel	Putative secreted protein
CG6234	Novel	Putative secreted protein, Glu-rich domain
CG10497	*Syndecan*, *sdc*	Transmembrane heparan sulfate
CG9623	*Inflated*, *if*	αPS2 integrin
CG6868	*tolloid*, *tld*	Metalloprotease
CG9458	Novel	GNS1/SUR4 membrane protein domain
CG15095	l(2)08717	General substrate transporter
CG5847	Novel	Endoglin/CD105 domain
CG12730	Novel	Membrane-associating domain (MARVEL)
CG3502	Novel	Membrane alanine aminopeptidase, Zn-binding site
CG15658	Novel	Leucine-rich repeat
CG10035	Novel	Transmembrane domain, coiled coil domain
Cytoskeletal and signal transduction proteins		
CG9366	*Rho-like*, *RhoL*	Rho Small monomeric GTPase
CG18485	*stumps*	FGF-receptor signaling pathway
CG10933	Novel	SH3 domains
CG15235	Novel	Rap guanyl-nucleotide exchange factor
CG12489	Novel	Band 4.1domain, RING-finger domain
CG11105	Novel	Calcium binding, EF-Hand family
CG5890	Novel	Calcium binding, EF-Hand family
Enzymes		
CG18516	*DmAO4*	Aldeyde oxidase
CG2915	Novel	Carboxypeptidase A
CG6805	Novel	Endonuclease/exonuclease/phosphatase family
Miscellaneous proteins		
CG12276	*Aos1*	Smt3-activating enzyme
CG1866	*Moca-cyp*	Nuclear cyclophilin
CG14112	*Soxneuro cofactor, SNCF*	Transcriptional activation
CG8625	*Imitation swi1*, *Iswi*	ATP-dependent nucleosome remodelling factor
CG11814	Novel	Beige/BEACH domain, WD-40 repeats
CG31163	Novel	SH3/SH2 domain, SAM domain
CG7956	Novel	Sac1 homology domain
CG11825	Novel	Hypoxia induced protein conserved region
Unknown		
CG13300	Novel	
CG40315	Novel	
CG5138	*Ocho*	
CG6478	Novel	Domain of unknown function (DUF243)
CG13023	Novel	
CG1674	Novel	
CG14803	Novel	
CG15735	Novel	
CG14112	Novel	Domain of unknown function (DUF733)

lated transcription factors (*twist*, *odd-skipped* and *sloppy-paired*) with known expression patterns and roles in early embryonic development (Thisse et al., 1988; Coulter et al., 1990; Grossniklaus et al., 1992). In addition, we isolated several novel genes that encode proteins with predicted RNA or DNA binding domains.

Among the genes coding for secreted and transmembrane proteins, we found *dilp4* that encodes an insulin-like peptide, and it has been implicated in cell growth and proliferation. This gene is expressed in the presumptive mesoderm and anterior midgut of early embryos (Casal and Leptin, 1996). The gene *inflated* encodes the α subunit of PS2 integrin, and loss-of-function mutations in this gene result in cell adhesion abnormalities in the gastrula (Roote and Zusman, 1995). Additional known genes were *Ance*, whose transcription takes place at the dorsalmost cells of embryos from stage 5 through stage 13 (Tatei et al., 1995), and *tolloid*, which encodes a secreted metalloprotease required to degrade the Short-gastrulation protein in the dorsal half of the embryo (Marques et al., 1997). Several novel genes that encode proteins with predicted signal peptide sequences and transmembrane domains were also isolated in our screening; among them, CG13333 was previously identified in a screening for neuronal precursor genes (Brody et al., 2002), and CG10035 was isolated by Kopczynski et al. (1998) in a screening for secreted and transmembrane proteins in *D. melanogaster* embryos (Clone ID: CK02586, Genebank accession AA141845). CG10035 corresponded to one of the nine transcripts identified in Kopczynski et al.'s screening as expressed during the gastrula stage.

Components of intracellular signaling pathways, such as *stumps*, an element of the FGF-receptor signaling pathway, and *RhoL*, a member of the Rho small GTPase family, were previously isolated from a subtractive cDNA library enriched in sequences expressed in the mesoderm (Casal and Leptin, 1996). In addition, we isolated a novel gene encoding a protein with a predicted band 4.1 domain, which is found in several cytoskeletal-associated proteins (Takeuchi et al., 1994), and CG15235, which is a cAMP-dependent guanine-nucleotide exchange factor of Rap1, a Ras-related GTPase involved in morphogenetic events during *D. melanogaster* gastrulation (Asha et al., 1999).

We compared the genes identified in this study with the freely available data (Yale Developmental Gene Expression database at http://genome.med.yale.edu/ Lifecycle) from a recent report on global gene expression analyses during *D. melanogaster* development (Arbeitman et al., 2002). Although the differences in the procedure to stage the embryos and in the hybridization techniques, we found that 11 of the genes reported here were also identified by Arbeitman et al. (2002) as early zygotic genes, including previously characterized (*twi*, *odd*, *stumps, sdc* and *Mocacyp*) and novel genes (CG7357, CG15095, CG15658 and CG7956). Arbeitman et al. (2002) detected high levels of expression of transient early zygotic genes only during the cellularization stage for a small number of genes (21 out of 534) possibly due to the fact that their microarray analysis were conducted with arrays that contained expression sequence tags (ESTs) that covered 30% of the *D. melanogaster* genome. The sensitivity of our approach based on a subtractive hybridization probably allowed us to detect additional zygotic genes whose level of expression transiently peaks during the gastrulation stage (Table 2).

We verified by semiquantitative PCR assay, the expression pattern for 24 randomly selected genes from our screening. In doing so, we simultaneously amplified the transcripts of selected genes and that of tubulin using cDNA from embryos at different stages of development (Table 2, Fig. 1). The relative expression level of each gene in the different stages was expressed as a percentage of its highest value at a given stage. The percentages were represented by a scale of intensity ranking from − to +++ (Table 2). Representative patterns of expression of four genes with their assigned scale of intensity are shown in Fig. 1. The results of the semiquantitative PCR assays indicated that all of the selected genes showed peaks of expression after stages 2–4 of embryogenesis. We found that 10 out of 24 genes showed peaks of expression in stages 6–7 (for instance, Fig. 1A, C and D), while 8 were expressed at their highest levels in stages 6–7 as well as in other stages

Table 2
Semiquantitative PCR analysis of selected genes

CG number	Relative expression levels[a]				
	2–4[b]	5[b]	6–7[b]	8–9[b]	10–12[b]
CG2915	–	++	+++	–	–
CG17829	–	+++	+++	–	–
CG15095	–	–	+++	+++	+++
CG7956	+	+++	+++	++	+
CG1866	–	++	+++	+++	++
CG18485 (*stumps*)	–	+++	+++	++	++
CG15658	–	++	+++	–	++
CG14803	+	++	+++	++	–
CG3851 (*odd*)	–	+	+++	++	++
CG13333	–	++	+++	++	+
CG11825	–	++	+++	+	–
CG12730	–	+++	+++	+++	–
CG10035	–	++	+++	++	–
CG13427	–	++	+++	+	–
CG8625 (*Iswi*)	–	–	+++	+	–
CG14112	–	+	+++	–	–
CG3502	–	+	+++	+++	++
CG17765	–	+++	+++	+++	+
CG31163	–	–	++	–	+++
CG8827 (*Ance*)	–	+	++	++	+++
CG11814	–	+	++	+++	+
CG15735	+	++	++	++	+++
CG9366 (*RhoL*)	–	+	++	+++	+++
CG12489	–	+	++	+++	–

[a] For each gene, 100% corresponds to its highest expression level in the different stages analyzed. Relative expression levels between: 0–20%, 21–50%, 51–80% and 81–100% are indicated as –, +, ++ and +++, respectively.

[b] Embryonic developmental stages.

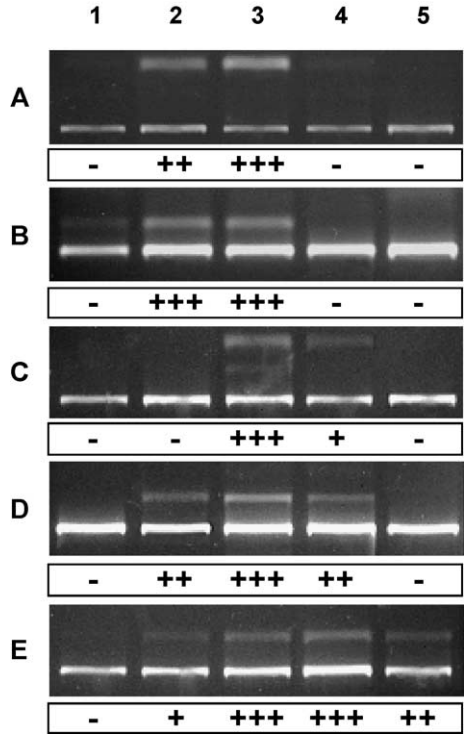

Fig. 1. Temporal expression of novel genes during *D. melanogaster* development. The transcripts of four genes isolated using a subtractive hybridization procedure were amplified, and their relative abundances were determined utilizing tubulin as an internal control. The amplification products of selected genes (upper band) and tubulin (lower band) are shown on the ethidium-bromide-stained 1.2% agarose gels. Developmental stages were the following: 2–4 (lane 1), 5 (lane 2), 6–7 (lane 3), 8–9 (lane 4) and 10–12 (lane 5). A: CG2815; B: CG17829; C: CG8625; D: CG10035; E: CG3502. A scale of relative intensity from − to +++ was assigned to the amplified products of the selected genes (see Table 2).

(for instance, Fig. 1B and D). Six genes were expressed at their highest levels in stages 8–9 (n=2), 10–13 (n=3) or both (n=1).

Whole-mount in situ hybridization analyses were performed for 20 of the 49 nonredundant genes to determine their spatial patterns of expression in developing embryos. We found that the expression of 65% of them was restricted to specific regions of the gastrula. The analysis of three of these genes that were highly expressed during gastrulation stage is described below.

3.3. Molecular characterization of three genes highly expressed in the gastrula

We further analyzed cDNA fragments that exhibited the highest levels of expression in embryos of stages 6–7. By searching the *D. melanogaster* databases made available by Berkeley *Drosophila* Genome Project (BDGP) (FlyBase, http://flybase.net/blast/), we found perfect matches of two cDNA inserts (255 and 413 bp) with the predicted gene CG13427 (318 bp) and other two inserts (625 and 989 bp) with the transcripts of gene CG6234 (2465 bp). An additional fragment of 450 bp was identical to the transcript (697 bp) of the previously reported gene *sala* (Reuter et al., 1996).

3.3.1. CG13427

The genome annotation from BDGP indicates that CG13427 is an intronless gene with no 5′-untranslated region. There are no expressed-sequence tags (ESTs) that match this gene, except for a cDNA sequence (Genebank Accession BQ102969) containing the entire open reading frame of CG13427. This sequence that was isolated in a screening for transcripts with restricted expression patterns in early embryos (Simin et al., 2002), along with one of the CG13427 clones isolated during our screening contained the 3′-untranslated region of this gene. To further characterize this novel gene, we decided to identify the full-length sequence of CG13427 using 5′-RACE–PCR. The primary PCR reaction on the embryonic total RNA templates did not produce a visible product (Fig. 2A, lane 2). A secondary PCR, using these reactions as templates, and nested CG13427 primers plus the SMART PCR primer did generate a product of approximately 400 bp and a smaller fragment of 200 bp (Fig. 2A, lane 3). The products were cloned and sequenced in both directions, revealing that the 400-bp product contained the CG13427 coding sequence, while the smaller product was a PCR artifact as it was not reproducible and could not be sequenced with the CG13427 primers used in the amplification reaction. When the sequences of the PCR products were matched to the known *D. melanogaster* genomic sequence, a 49-bp 5′-untranslated region was identified containing a perfect consensus for the Initiator (Inr) TCAGTT, beginning with the A at +1 (Fig. 2A). There are no consensus sequences for TATA box elements, except for a partial match (TATTTAA) located 33 bp upstream of the transcription start site. Instead, there is a probable downstream promoter element (DPE) at sites +28–33 (ACACGC) that match the DPE consensus sequence (Kutach and Kadonaga, 2000). Our results on CG13427 allowed us to define the transcriptional initiation for this gene that has not been predicted before. CG13427 promoter does not contain a prototypical TATA box sequence and is likely to be DPE-dependent. The DPE functions cooperatively with an Inr sequence at strictly maintained positions with respect to the transcriptional initiation site to bind to $TAF_{II}60$ and direct accurate and efficient initiation of transcription in TATA-less promoters (Burke and Kadonaga, 1997).

The presence of a polyadenylation signal (AATAAA) 24 bp from the poly(A) site along with our results on the 5′-RACE analysis allowed us to predict a size of 450 bp for the full-length cDNA of CG13427. The conceptual translation of CG13427 from the first ATG indicates that it codes for a 105-amino-acid (aa; 11.7 kDa) polypeptide, with no significant similarity with any other characterized protein from *D. melanogaster* or other species. The protein coded by CG13427 contains a stretch of hydrophobic residues near the N-terminus, which is similar to a signal sequence

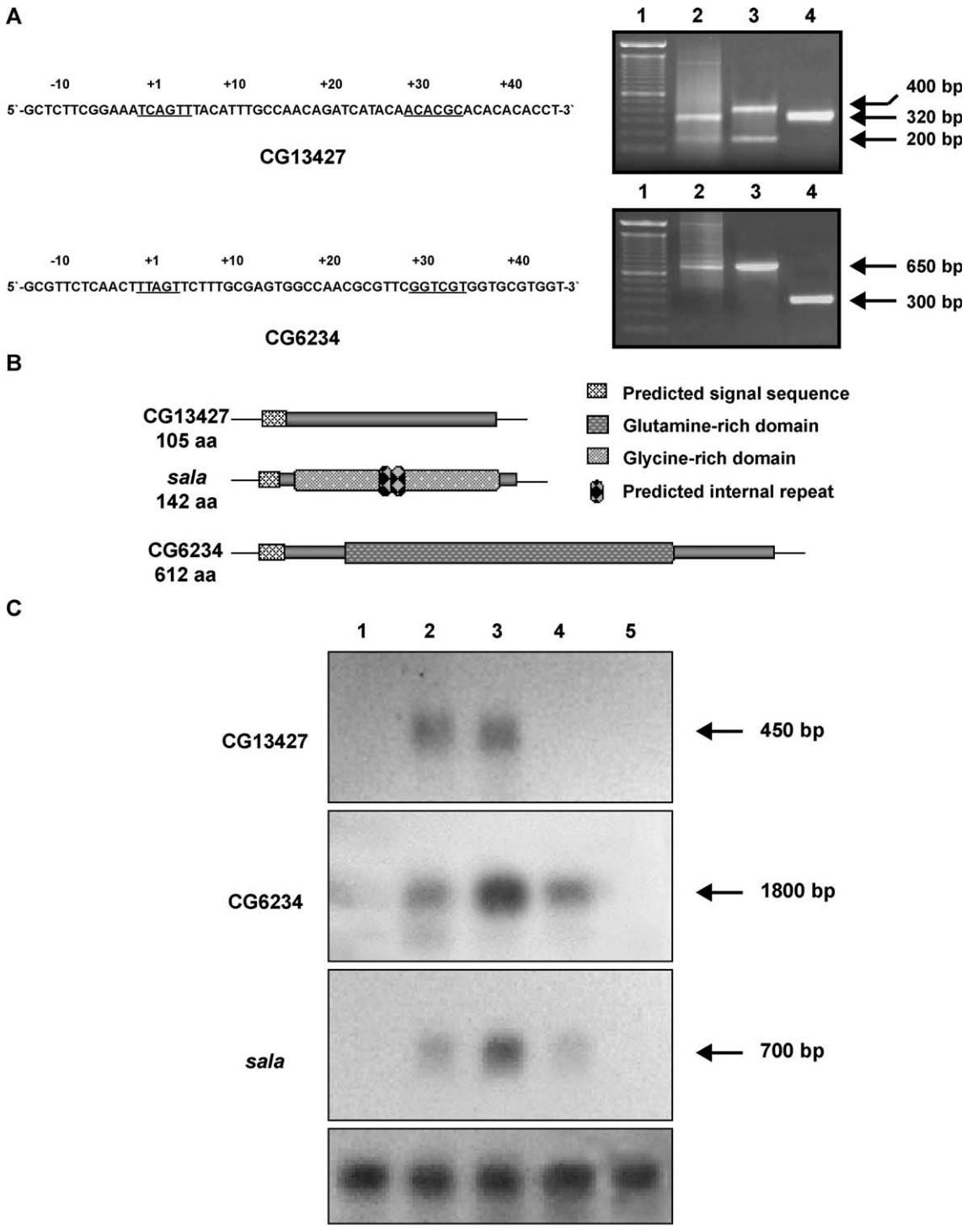

Fig. 2. Molecular characterization of CG13427, CG6234 and *sala*. (A) Transcriptional start and 5′-untranslated sequence. PCR products generated from 5′-RACE analysis of CG13427 and CG6234 are shown on the 1% agarose gel. The markers are DNA segments increasing at 100-bp increments (lane 1). Lane 2, primary PCR reaction on embryonic total RNA templates; lane 3, secondary PCR using primary PCR reactions as templates; lane 4, control PCR reactions using gene-specific sense and antisense primers. The sequences on the side of each gel show the nucleotide sequence of the 5′-RACE product aligned with the genomic DNA sequence from CG13427 and CG6234 gene regions. For each gene, the Inr site is underlined and begins at −2. The DPE site is also underlined and begins at position −28. (B) The schematic shows the predicted structure of the proteins encoded by the three genes. (C) Developmental Northern blot analyses. Total RNA (8 μg) from stages 2–3 (lane 1), 5 (lane 2), 6–7 (lane 3), 8–9 (lane 4) and 10–13 (lane 5) were hybridized with ^{32}P-labeled probes for CG13427 (450 bp), CG6234 (637 bp), *sala* (420 bp) and β-tubulin (460 bp). Expression of tubulin was used as a loading control to calculate the relative abundance of the transcripts. Estimated molecular weight of mRNA is marked on the right side of the picture. Northern blot analyses were performed as described in Section 2.4 using 1.2% formaldehyde-agarose gels.

commonly seen in secreted proteins. This is supported by the finding of a putative cleavage site that follows the signal sequence in position 19–20 (Fig. 2B). Analysis of its sequence by PSORT I and II (http://psort.nibb.ac.jp/) indicate that CG13427 has a high probability of being secreted. However, the genome-wide analysis of protein interaction map of *D. melanogaster* (Fly GRID Interaction Data, http://biodata.mshri.on.ca/fly_grid/servlet/SearchPage) inferred interactions between CG13427 and two cytoplasmic proteins based on yeast two-hybrid experiments (Giot et al., 2003). The lack of consistency between a putative extracellular localization of CG13427 and these reported interactions will need to be solved by directly addressing whether CG6234 is secreted by the cells and by subcellular distribution analyses of this gene.

For Northern blot analysis, total RNA (8 μg) was isolated from embryos at stages 2–3, 5, 6–7, 8–9 and 10–13. A 255-bp cDNA fragment of gene CG13427 was used as a probe revealing a band of approximately 450 bp that matched the predicted size for CG13427 mRNA. Taken together, the characterization of CG13427 indicated that it is an intronless gene with transient expression in stages 5 and 6–7 of embryogenesis (Fig. 2C, lanes 2–3). Although, in higher eukaryotes, most genes contain introns, there are 1288 intronless genes (approximately 10% of the protein coding genes) in *D. melanogaster* (http://sege.ntu.edu.sg/wester/intronless/index.html). One known family of intronless genes encodes *D. melanogaster* heat shock proteins (Michaud et al., 1997). Another family encodes glutathione transferases (Toung et al., 1993). It has been argued that, in the absence of posttranscriptional splicing, intronless genes may be expressed more efficiently in response to extracellular signals (Zhou and Chiang, 2001). Further analysis of the promoter region of CG13427 will be required to determine the transcriptional regulation of this gene.

3.3.2. CG6234

ICG6234 showed no homology to any known gene but matched a cluster of ESTs. The expression pattern of CG6234 during *D. melanogaster* development is reported in the Yale Developmental Gene Expression database. High relative levels of CG6234 expression during embryonic stages 5 to 16 were determined from microarray hybridization data (Arbeitman et al., 2002). Using Northern blot assays, we detected expression of CG6234 only during stages 5 to 10 of development. Similarly, our in situ hybridization analysis showed that the expression of this gene was undistinguishable after stage 11 (see below). In the same database, additional high levels of CG6234 expression were reported during prepupal and pupal stages; this is in agreement with the results of Butler et al. (2003) who showed a 2.8-fold enrichment of CG6234 transcripts in the wing imaginal disc (Table 2 in Butler et al., 2003). These data, along with our results that show high levels of expression of CG6234 during gastrulation, suggest that this gene is transcriptionally regulated during the life cycle of *D. melanogaster*, reaching its highest levels of expression during embryogenesis and metamorphosis, the two major periods of morphological changes (Reinke and White, 2003).

To identify the full-length cDNA of CG6234, we applied the 5′-RACE–PCR analysis as described above. The secondary PCR reaction permitted to amplify a single product of approximately 650 bp (Fig. 2A, lane 3). A 299-bp 5′-untranslated region was identified after matching the sequence of the PCR product to *D. melanogaster* genomic sequence. As in the case of CG13427, the gene CG5234 lacks a TATA box consensus sequence, and its Inr sequence (TTAGTT) does not perfectly match the consensus sequence. However, inspection of the *D. melanogaster* core promoter database (http://www-biology.ucsd.edu/labs/Kadonaga/DCPD.html) that comprise 205 sequences indicates the occurrence of T at position -1, suggesting that this sequence falls within the range of sequences that have been shown to function as Inr (Kutach and Kadonaga, 2000). In addition, a good match with the consensus DPE sequence was identified at site +28–33 (GGTCGT) of gene CG6234 (Fig. 2A). Since both of our clones contained the 3′-untranslated region of this gene, we were able to identify the presence of a polyadenylation signal (AATAAA), 17 bp from the poly(A) site. Therefore, after 5′-RACE–PCR, a full-length cDNA of 1838 bp was determined, which matched the cDNA sequence reported by BDGP. Our analysis of CG6234 indicates that the previously published cDNA sequence is complete and confirms the position of transcriptional initiation predicted by the *D. melanogaster* genome project.

The genome annotation from BDGP indicates that CG6234 is 2465 bp in length and that it has three exons (GenBank accession AY052038). The 1838-bp cDNA sequence of CG6234 encodes a polypeptide of 612 amino acids that contains a Glu-rich region extending from amino acids 124 to 489 and a signal peptide sequence (amino acids 1 to 40) followed by a putative cleavage site (Fig. 2B). CG6234 showed low similarity to mouse Zonadhesin precursor; this homology, however, entirely resides in the Glu-rich domain of CG6234 polypeptide. Northern blot hybridizations using a 637-bp fragment of CG6234 revealed a single mRNA of the predicted size that showed a peak of expression in embryos of stages 6–7 (Fig. 2C, lanes 2–4).

3.3.3. sala

sala transcript can be translated into a short polypeptide composed of 142 amino acids, containing a series of Gly–Gly–X repeats with low similarity to several collagen proteins. The *sala* protein contains a stretch of 11 amino acids in its N-terminal that seems to function as a signal sequence (Fig. 2B). Reuter et al. (1996) showed that the protein is predominantly secreted from culture cells after transfection with *sala* expression vectors. Moreover, the expression of a modified Sala lacking 9 amino acids of the

putative signal sequence indicates that this sequence is required for the secretion of the processed protein in culture cells. However, inspection of the Fly GRID Interaction Data revealed that a cytoplasmic localization of *sala* could be inferred based on the reported interactions between *sala* and two well-known cytoplasmic proteins (Giot et al., 2003). Further biochemical studies of these potential interactions of Sala would help clarify the discrepancy between a cytoplasmic or extracellular localization of this protein. The function of *sala* is unknown, and ectopic expression or absence of *sala* protein had no effect on embryonic development, suggesting that *sala* has redundant roles during development (Reuter et al., 1996).

In addition to the previously reported expression pattern of *sala* (Reuter et al., 1996), our Northern blot analysis revealed a single 700-bp transcript whose expression level increased by more than fivefold in stages 6–7 of development (Fig. 2C, lanes 2–3). The expression patterns of this gene indicate that it is transiently expressed during early stages of embryogenesis, reaching their highest level of expression during gastrulation stage.

3.4. Spatial expression of sala, CG13427 and CG6234 during early stages of development as revealed by in situ hybridizations

The spatial expression patterns of CG13427, CG6234 and *sala* were studied by whole-mount in situ hybridizations of early embryos (Fig. 3). A probe prepared from our CG13427 clone showed that the expression of this gene was undetectable in the syncytial blastoderm stage (A). In the cellular blastoderm, CG13427 was expressed in the dorsolateral region of the embryo, with some staining reaching the ventral regions at 60–70% and 15–45% egg

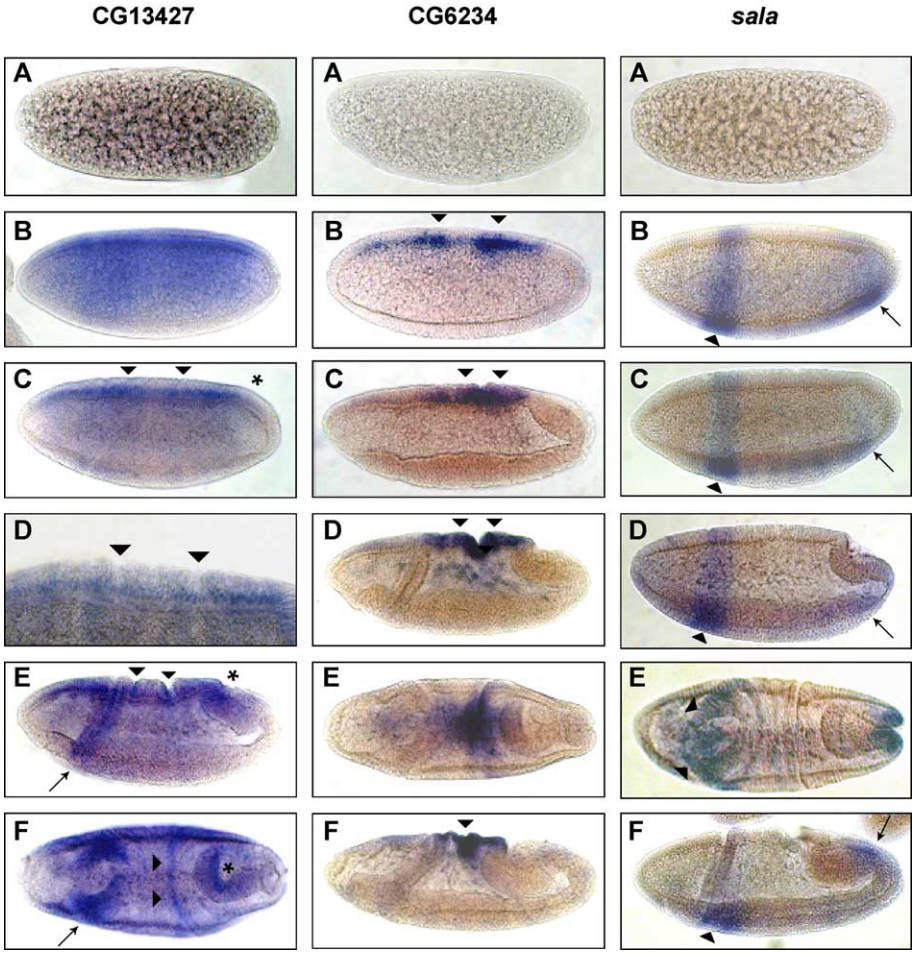

Fig. 3. Distribution of CG13427, CG6234 and *sala* transcripts during *D. melanogaster* embryogenesis. Whole-mount in situ hybridization of embryos using DIG-labeled DNA probes was carried out as described in Section 2.7. Embryos are oriented with the anterior region to the left and the dorsal region facing upward; all are lateral views unless stated otherwise. Panels (A) and (B) correspond to syncytial blastoderm (stage 3) and cellular blastoderm (stage 5), respectively. CG13427: (C–D) gastrula (stages 6); (E–F) gastrula (stage 7). Transcripts are expressed in the dorsal cells, including those forming the dorsal transverse furrows (arrowheads), the cephalic furrow (arrow) and the proctodeal primordium (asterisks); (D) detail showing dorsal transverse furrows; (F) dorsal view. CG6234: (C) gastrula (stage 6); (D–E) gastrula (stage 7), expression in the dorsal cells is shown (arrowheads); (E) dorsal view. (F) stage 8, transcripts are expressed in the amnioserosa cells. *sala*: (C) gastrula (stage 6); (D–E) gastrula (stage 7) showing expression in the cephalic furrow (arrowheads) and in cells located at the posterior half of the proctodeal primordium (arrow); (F) stage 8; (E) dorsal view.

length (B). At the beginning of gastrulation (C), the expression of CG13427 became more restricted to the elongate dorsal cells of the embryo, including the dorsal transverse furrows (C–E, arrowheads). During gastrulation (E–F), expression of CG13427 was also observed in the primordium of the proctodeum (asterisk), as well as in the cells that form the cephalic furrow (arrow). In the case of CG6234, no maternal accumulation was observed in the syncytial blastoderm stage (A). Spatial expression of CG6234 was highly restricted during the stages analyzed. It was first detected at stage 5 in the dorsalmost cells (B, arrowheads), and this expression pattern persisted in these cells during gastrulation stages 6 (C) and 7 (D–E) and during stages 8–9 when they differentiate into the amnioserosa (F, arrowhead). The expression of CG6234 in the cells that formed the amnioserosa persisted during the period of germ band extension and became undetectable after stage 11. Concerning the spatial distribution of *sala* mRNA, it was initially detected in the cellular blastoderm embryo in an anterior region at 60–70% egg length, where the cephalic furrow will form at the beginning of gastrulation (B, arrowhead). During this stage of development, *sala* was also expressed in a region at 15–20% egg length (arrow), showing a weaker staining along the midventral domain (B). During stages 6 to 8, strong *sala* expression was detected in the cells that form the cephalic furrow and the ventroposterior region of this invagination (C–E, arrowheads). At stage 8, the posterior domain of *sala* expression became restricted to a group of cells located at the caudal side of the proctodeal primordium (F, arrow).

The spatial expression patterns of CG13427 and *sala* in the invaginating cells located on the cephalic furrow, the dorsal transverse furrows and the proctodeal primordium suggests a possible involvement of CG13427 and *sala* in these gastrulation events. The behavior of the cells at these regions seems to be different (Costa et al., 1993). The cephalic and dorsal transverse furrow formation implicates nuclear migration and cell shortening along the apical–basal axis, while the proctodeal invagination appears to be driven by similar cellular mechanisms as the ventral furrow, involving cell surface flattening and apical constriction. Despite morphological differences, all these cellular events are involved in a common process—the invagination of the blastoderm epithelium (Schöck and Perrimon, 2002). CG6234, on the other hand, is expressed in the amnioserosa, an extraembryonic membrane that arises from a group of dorsal epithelial cells in the stage 5 embryo. These cells undergo a transformation in shape from columnar to squamous to become the sheet of thin cells that form the amnioserosa at stages 8–9 (Campos-Ortega and Hartenstein, 1985). The expression pattern of CG6234 during stages 5 to 9 of embryogenesis suggests a possible involvement of this gene in the early events of amnioserosa formation.

The three genes that we described here met the aims of our analysis. They are novel genes expressed at their highest level during the gastrulation stage. Moreover, the transcripts of these genes are present in the cells that undergo morphological changes during this stage of development, suggesting their involvement in cell shape changes and movements that drive the remodeling of the embryonic epithelium during the gastrulation stage (Schöck and Perrimon, 2002). Sequence analysis of the proteins encoded by these genes indicates the presence of consensus signal sequences, suggesting that they might be putative secreted proteins; for one of them, previous evidence supports this possibility. The idea that the three genes characterized here, CG13427, CG6234 and *sala*, encode putative secreted proteins is suggestive since secreted and transmembrane proteins are expected to be active at and play an important role in the temporal and spatial coordination of epithelium morphogenesis during gastrulation.

4. Conclusions

(1) The screening described here allowed us to identify novel genes differentially expressed between gastrulation and syncytial blastoderm stages and confirmed the expression patterns of several already known genes. Sixty-five percent of the identified sequences corresponded to uncharacterized genes (n=33), indicating that the SSH procedure together with the precise selection of staged embryos may be useful to provide information on novel genes with potential function at specific stages of development.

(2) Temporally and spatially restricted patterns of expression in developing embryos were observed for the three genes characterized here. The transcripts of two of them were detected in groups of cells that invaginate during gastrulation, while the third was expressed in a subset of cells that differentiate into the amnioserosa tissue. The expression pattern of these genes during *D. melanogaster* embryogenesis suggests their possible involvement in the remodeling of the epithelium formed during embryo cellularization.

(3) Since two of these genes, CG13427 and CG6234, are novel sequences, we define their transcription initiation site and 5′-untranslated region. The results showed that they do not contain a prototypical TATA box sequence and that they are probably DPE-dependent for their transcription. Further analysis of the promoter region of CG13427 and CG6234 will provide insights into their transcriptional regulation during *D. melanogaster* development.

Acknowledgments

This work was supported by Fondecyt 1010693 and 1030618.

References

Arbeitman, M.N., et al., 2002. Gene expression during the life cycle of *Drosophila melanogaster*. Science 297, 2270–2275.

Asha, H., de Ruiter, N.D., Wang, M.-G., Hariharan, I.K., 1999. The Rap1 GTPase functions as a regulator of morphogenesis in vivo. EMBO J. 18, 605–615.

Ausubel, F., et al., 1997. Short Protocols in Molecular Biology. John Wiley & Sons.

Brody, T., Stivers, C., Nagle, J., Odenwald, W.F., 2002. Identification of novel *Drosophila* neural precursor genes using a differential embryonic head cDNA screen. Mech. Dev. 113, 41–59.

Burke, T.W., Kadonaga, J.T., 1997. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. Genes Dev. 11, 3020–3031.

Butler, M.J., et al., 2003. Discovery of genes with highly restricted expression patterns in the *Drosophila* wing disc using DNA oligonucleotide microarrays. Development 130, 659–670.

Campos-Ortega, J.A., Hartenstein, V., 1985. Stages of *Drosophila* embryogenesis. In: Campos-Ortega, J.A., Hartenstein, V. (Eds.), The Embryonic Development of *Drosophila melanogaster*. Springer-Verlag New York, pp. 9–84.

Casal, J., Leptin, M., 1996. Identification of novel genes in *Drosophila* reveals the complex regulation of early gene activity in the mesoderm. Proc. Natl. Acad. Sci. U. S. A. 93, 10327–10332.

Costa, M., Sweeton, D., Wieschaus, E., 1993. Gastrulation in *Drosophila*. In: Bate, M., Martinez Arias, A. (Eds.), The Development of *Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, pp. 425–465.

Costa, M., Wilson, E., Wieschaus, E., 1994. A putative cell signal encoded by the folded gastrulation gene coordinates cell shape changes during *Drosophila* gastrulation. Cell 76, 1075–1089.

Coulter, D.E., Swaykus, E.A., Beran-Koehn, M.A., Goldberg, D., Wieschaus, E., Schedl, P., 1990. Molecular analysis of odd-skipped, a zinc finger encoding segmentation gene with a novel pair-rule expression pattern. EMBO J. 9, 3795–3804.

Diatchenko, L., et al., 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. U. S. A. 93, 6025–6030.

Giot, L., et al., 2003. A protein interaction map of *Drosophila melanogaster*. Science 302, 1727–1736.

Gong, B., Ge, R., 2000. Using the SMART cDNA system to map the transcription initiation site. BioTechniques 28, 846–851.

Grossniklaus, U., Pearson, R.K., Gehring, W.J., 1992. The *Drosophila* sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. Genes Dev. 6, 1030–1051.

Kopczynski, C.C., et al., 1998. A high throughput screen to identify secreted and transmembrane proteins involved in *Drosophila* embryogenesis. Proc. Natl. Acad. Sci. U. S. A. 95, 9973–9978.

Kutach, A.K., Kadonaga, J.T., 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. Mol. Cell. Biol. 20, 4754–4764.

Lehmann, R., Tautz, D., 1994. In situ hybridization to RNA. In: Goldstein, L.S.B., Fyrberg, E.A. (Eds.), Methods in Cell Biology. Academic Press, pp. 575–598.

Leptin, M., 1999. Gastrulation in *Drosophila*: the logic and the cellular mechanisms. EMBO J. 18, 3187–3192.

Marques, G., Musacchio, M., Shimell, M.J., Wunnenberg-Stapleton, K., Cho, K.W., O'Connor, M.B., 1997. Production of a DPP activity gradient in the early *Drosophila* embryo through the opposing actions of the SOG and TLD proteins. Cell 91, 417–426.

Michaud, S., Marin, R., Tanguay, R.M., 1997. Regulation of heat shock gene induction and expression during *Drosophila* development. Cell. Mol. Life Sci. 53, 104–113.

Miklos, G.L., Rubin, G.M., 1996. The role of the genome project in determining gene function: insights from model organisms. Cell 86, 521–529.

Reinke, V., White, K.P., 2003. Developmental genomic approaches in model organisms. Annu. Rev. Genomics Hum. Genet. 3, 153–178.

Reuter, D., et al., 1996. Regulation, function and potential origin of the *Drosophila* gene spalt adjacent, which encodes a secreted protein expressed in the early embryo. Chromosoma 104, 445–454.

Roote, C.E., Zusman, S., 1995. Functions for PS integrins in tissue adhesion, migration, and shape changes during early embryonic development in *Drosophila*. Dev. Biol. 169, 322–336.

Schöck, F., Perrimon, N., 2002. Molecular mechanisms of epithelial morphogenesis. Annu. Rev. Cell Dev. Biol. 18, 463–493.

Simin, K., et al., 2002. Profiling patterned transcripts in *Drosophila* embryos. Genome Res. 12, 1040–1047.

Takeuchi, K., Kawashima, A., Nagafuchi, A., Tsukita, S., 1994. Structural diversity of band 4.1 superfamily members. J. Cell. Sci. 107, 1921–1928.

Tapia, L., et al., 2004. Metallothionein is key for safe intracellular copper storage and cell survival at normal and supra-physiological exposure levels. Biochem. J. 378, 617–624.

Tatei, K., Cai, H., Ip, Y.T., Levine, M., 1995. Race: a *Drosophila* homologue of the angiotensin converting enzyme. Mech. Dev. 51, 157–168.

Taylor, M.V., 2000. A novel *Drosophila* mef2-regulated muscle gene isolated in a subtractive hybridization-based molecular screen using small amounts of zygotic mutant RNA. Dev. Biol. 220, 37–52.

Thisse, B., Stoetzel, C., Gorostiza-Thisse, C., Perrin-Schmitt, F., 1988. Sequence of the twist gene and nuclear localization of its protein in endo-mesodermal cells of early *Drosophila* embryos. EMBO J. 7, 2175–2183.

Toung, Y.P., Hsieh, T.S., Tu, C.P., 1993. The glutathione S-transferase D genes. A divergently organized, intronless gene family in *Drosophila melanogaster*. J. Biol. Chem. 268, 9737–9746.

Wieschaus, E., 1996. From molecular patterns to morphogenesis: the lessons from *Drosophila*. In: Frängsmyr, T. (Ed.), Les Prix Nobel (The Nobel Prizes) 1995. Almqvist and Wiksell International, Stockholm, Sweden, pp. 235–260.

Wieschaus, E., Sweeton, D., 1988. Requirements for X-linked zygotic gene activity during cellularization of early *Drosophila* embryos. Development 104, 483–493.

Zhou, T., Chiang, C.M., 2001. The intronless and TATA-less human TAF(II)55 gene contains a functional initiator and a downstream promoter element. J. Biol. Chem. 276, 25503–25511.