Genetics of arrhenotokous and thelytokous reproduction in Venturia canescens (Hymenoptera)

Mateo Leach, Irene Victoria

CHAPTER 6

Transcriptome and proteome analysis of ovaries of arrhenotokous and thelytokous *Venturia canescens*

Irene Mateo Leach
Anne Hesseling
Wim H.C. Huibers
Hanneke Witsenboer
Leo W. Beukeboom
Louis van de Zande

Abstract
Under arrhenotoky, unfertilized haploid eggs develop as males but under thelytoky they develop into diploid females after they have undergone diploidy restoration. In the parasitoid wasp *Venturia canescens* both reproductive modes occur. Thelytoky is genetically determined but the underlying genetics of diploidy restoration remains unknown. In this study we aim to identify the genes and/or proteins that control thelytoky in this wasp. cDNA-AFLP analysis of total ovarian RNA and two dimensional protein electrophoresis in combination with mass spectrometry revealed putative transcripts and proteins involved in arrhenotokous and thelytokous processes. The detected tubulin and actin protein differences are most likely functionally connected to the two types of parthenogenetic development.

Insect Molecular Biology (2009) 18: 477–482.

Introduction

All Hymenoptera have a haplodiploid mode of sex determination. Within haplodiploidy, arrhenotoky is the most common mode of reproduction: unfertilized eggs develop into males that are haploid, whereas fertilized eggs develop into diploid females. Thelytoky is less common, but nevertheless occurs in many hymenopteran groups. Thelytokous females develop parthenogenetically from unfertilized eggs after restoration or maintenance of diploidy and males do not occur. Hence, arrhenotokous males and thelytokous females are both produced parthenogenetically under haplodiploidy.

Sympatrical coexistence of arrhenotokous and thelytokous reproduction raises many evolutionary questions since asexual reproduction has a short term competitive advantage over sexual reproduction. Genes that confer parthenogenetic reproduction (thelytoky) are expected to rapidly spread in a sexual (arrhenotokous) population, because they are transmitted twice as efficiently as sexual genes (twofold cost of sex theory, Maynard Smith, 1978). On the other hand, thelytokous populations have the long term disadvantage of accumulating deleterious mutations (Muller, 1964; Kondrashov, 1982) and have a reduced evolutionary potential to adapt to changing environments in contrast to arrhenotokous populations (van Valen, 1973; Ghiselin, 1974).

Most known cases of thelytoky are caused by microorganisms (van Wilgenburg *et al.*, 2006). However, a genetic basis for thelytoky has been shown for some species of wasps, ants and the cape honey bee (Tucker, 1958; Stouthamer *et al.*, 1990; Belshaw *et al.*, 1999; Schilder *et al.*, 1999; Beukeboom & Pijnacker, 2000; Vavre *et al.*, 2004; Lattorff *et al.*, 2005). Recently, Lattorff *et al.* (2005) have found that thelytokous parthenogenesis in the cape honey bee is a qualitative character determined by a single recessive gene, called *thelytoky* (*th*). Many different forms of thelytoky are known to occur within the Hymenoptera based on cytogenetical mechanisms (van Wilgenburg *et al.*, 2006), but the molecular mechanisms underlying these forms of parthenogenesis are completely unknown.

Venturia canescens (Gravenhorst) is a solitary endoparasitoid wasp of lepidopteran larvae (Beling, 1932; Salt, 1976). It has been widely used as a biological model in behavioural, population dynamical, genetical and physiological studies (references in Beukeboom & Pijnacker, 2000; Thiel *et al.*, 2007). The reproductive mode is a polymorphism in the species that is fixed at the individual level; arrhenotokous females cannot switch to thelytoky or vice versa and females of either mode occur sympatrically in Southern Europe (Schneider *et al.*, 2002). Thelytokous females produce haploid eggs meiotically that subsequently undergo diploidy restoration and develop into diploid females (Speicher, 1937; Beukeboom & Pijnacker, 2000). Beukeboom and Pijnacker (2000) showed that thelytoky in *V. canescens* is not due to infection with *Wolbachia* and supposed a genetic basis for thelytoky. In *V. canescens* it is not known whether a single or multiple genes are

responsible for thelytoky. It can also not be excluded that thelytoky is encoded by mitochondrial genes that are maternally inherited.

We set out to perform crosses between arrhenotokous males and thelytokous females, which would allow us to study the genetic basis of thelytoky in *V. canescens*, but unfortunately, were not able to repeat the results of Schneider *et al.* (2003). The most likely explanation is that gene exchange between arrhenotokous males and thelytokous females is very rare. As an alternative approach to elucidate the genetic basis of thelytoky we compared transcripts and proteins in the ovaries of arrhenotokous and thelytokous females. We used ovaries because the signal for parthenogenetic male or female development appears to be already present in the egg upon oviposition. Diploidy restoration in the thelytokous strains of *V. canescens* is described as a form of central fusion automictic parthenogenesis (Beukeboom & Pijnacker, 2000). We know that thelytokous and arrhenotokous oogenesis follow different ways, but it is not known what triggers the thelytokous egg to become diploid when it is not fertilized, whereas an arrhenotokous egg remains haploid.

In this study, we aim to identify genes that are differentially regulated during oogenesis in arrhenotokous and thelytokous females of *V. canescens*. We compare the gene and protein expression patterns in the ovaries of unmated females of both reproductive modes. For transcriptome analysis we use the cDNA-AFLP® gene expression technique (Vos *et al.*, 1995; Bachem *et al.*, 1996; Vuylsteke *et al.*, 2007) and for the proteomics two-dimensional gel electrophoresis and mass spectrometry. This study provides the first candidate genes for thelytokous reproduction in *V. canescens*.

Materials and methods

RNA extraction

Two independent samples of 10 virgin females from arrhenotokous strains with four individuals from Mont Boron-Valbonne, two from Antibes and four from Valence were collected. Similarly, two independent samples were collected from thelytokous strains with two individuals from Mont Boron, two from Mont Boron 2003, two from Valbonne, two from Valence, one from Antibes and one from Antibes 2003. Ovaries were pooled per reproductive mode to ensure that any possible difference in the gene and protein expression between the two reproductive modes was not due to genetic variation between populations. Newly emerged females were kept individually for two days with honey to increase egg production before dissection. Ovaries were dissected from the abdomen in phosphate-buffered saline (PBS). Tissue was disrupted in 350 µl RLT lysis buffer (QIAGEN) and ß-mercaptoethanol using a T25 Basic Ultra-Turrax® homogenizator (IKA Labortechnik). The QIAprep Spin Miniprep Kit (QIAGEN) was used for total RNA

extraction following the manufacturer's protocol. After DNase treatment, quality and quantity of RNA was checked on a RNA denaturing gel and on a spectrophotometer (Nanodrop nd-1000).

cDNA-AFLP analysis and sequencing

Double-stranded cDNA was synthesized from 20 µg of total RNA for each one of the two arrhenotokous and two thelytokous independent samples and cDNA-AFLP analysis was further performed according to Vuylsteke *et al.* (2007). cDNA was digested with the 4-bp restriction endonucleases *TaqI* and *MseI* restriction enzymes. The choice of using these restriction enzymes was based on Reineke *et al.* (2003). These authors had previously tried several enzyme combinations for cutting *V. canescens* cDNA and found these two enzymes to be the best cutters for this species. After adaptor ligation (*MseI* 5′-GACGATGAGTCCTGAG-3′ and 3′-TACTCAGGACTCAT-5′; *TaqI* 5′-CTCGTAGACTGCGTAC-3′ and 3′-TCTGACG-CATGGC-5′), preamplification was done using the corresponding primer for the enzyme adaptor sequence without extension (5′-GTAGACTGCGTACCGA-3′ as *TaqI* primer and 5′-GATGAGTCCTGAGTAA-3′ as *MseI* primer). An aliquot of the preamplification products was selectively amplified using the preamplification primers with an extra 2 or 3 nucleotide extension at the 3′ end for the *TaqI* and *MseI* primer respectively. The cDNA-AFLP analysis was performed in duplo with two independent samples per reproductive mode which includes a biological control of the analysis and excludes any variation due to the technical procedure. Fragments were visualized using infrared dye 700 for product detection on an automated LI-COR system (Biolegio). A selection of fragments that were differentially expressed in the two samples were further sequenced. Similarity searches were done by comparing the cDNA nucleotide sequences against the reference protein database of *Apis mellifera* (genome built 4.0) and *Nasonia vitripennis* (genome built 1.1) using program BlastX and tBlastn from NCBI. Homologous sequences with an E-value lower than $1x10^{-6}$ were counted as significant (Table 6.1). These results were also analysed using homology search programs FASTA (Pearson *et al.*, 1997) and WU-Blast2 (www.ei.ac.uk/blast2/).

Protein extraction

A preliminary Bradford assay was done to measure total protein content in two arrhenotokous and two thelytokous wasps. The total protein content of individual ovaries was 0.02 µg/µl and similar for both reproductive modes, indicating that the same number of arrhenotokous and thelytokous ovaries could be pooled to collect the sample for the two-dimensional gel electrophoresis.

As for the RNA, samples were pooled from the ovaries of 180 females of each reproductive mode to exclude as much as possible false differences caused by variation between lines within each mode. Samples were dissected on PBS and collected in 500 µl sample buffer containing 9.5 M urea (Fluka), 2% 3-[(cholamodo-

propyl)-dimethylammonio]-1-propanesulfonate (CHAPS) (Sigma), 2 mM tributilphosphine (TBP) (Bio-Rad) and 0.8% carrier ampholites (Bio-lyte, Bio-Rad). No protease inhibitors were used neither in during dissection nor sample preparation. Using 500 mg glass beads, cells were disrupted in a Fast Prep F120 (Savant) and supernatant was collected by centrifugation

Protein electrophoresis

IsoElectric Focusing (IEF) was performed by loading 250 µl of sample onto a ReadyStrip IPG strip pH 4-7 (Bio-Rad) and rehydrated overnight (Rabilloud *et al.*, 1994). IEF was performed at 50 µA/strip, 150 V for 30 min, 300 V for 1 h, 600 V for 1 h, and 8000 V until steady state (22000 Vh) at 20°C.

For the second dimension electrophoresis (SDS-PAGE), the IPG strips were equilibrated in buffer (0.05 M Tris-HCL pH 8.8, 6 M urea, 30% w/v glycerol and 2% w/v SDS) with 1% w/v dithiothreitol (DTT) (Sigma) and 4% iodoacetamide (Sigma) according to Görg *et al.*, (2000). Equilibrated IPG strips were fixed to a precast 12.5% polyacrylamide gel (Criterion Tris-HCl gel 13.3 x 8.7 cm, Bio-Rad) with a 0.5% agarose gel solution containing 25 mM Tris, 192 mM glycine, 0.1% SDS and a trace of bromophenol blue. We used the PageRulerTM Prestained Protein Ladder (Fermentas) as reference marker. Electrophoresis was performed at room temperature and 100 V for 2.5 h. Gels were fixed with a solution of 40% methanol and 10% acetic acid for 30 min followed by Coomassie Brilliant Blue staining (Page Blue, Fermentas) at room temperature for 3 h and subsequent overnight destaining in milliQ water. Gels were analyzed and scanned using the PD*Quest* Image analysis software (Bio-Rad). Differentially expressed spots were selected on basis of absence/presence and excised from the gel for in gel digestion and MALDI (matrix assisted laser desorption ionization) mass spectrometry analysis.

Mass Spectrometry analysis

The gel slices were washed with 50% acetonitrile in 25 mM NH_4HCO_3 until they were completely destained. After vacuum drying, gel slices were rehydrated in 5 µl trypsin solution (Promega V5111; 12.5 ng/µl in 25 mM NH_4HCO_3), with an overlay of 20 µl of 25 mM NH_4HCO_3 and incubated overnight at 37°C. To recover the tryptic peptides the gel slices were eluted three times with 20 µl 75% acetonitrile and 0.1% aqueous trifluoric acid (TFA). The extracted peptides were concentrated under vacuum. Finally, the dry peptides were dissolved in 10 µl 0.1% TFA in 5% acetonitrile. Mass spectrometric analysis was performed using a MALDI-TOF/TOF 4700 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA).

Protein identification was carried out using Mascot, local host (www.matrixscience.com), searching against the *Apis mellifera* and *Nasonia vitripennis* protein sequence databases at the National Centre for Biotechnology Information (NCBI). Peptide tolerance was set to 50 ppm and MS/MS tolerance to 0.2 Da, allowing for 2 missed trypsin cleavages and variable modification for oxidation of methionine.

Results with at least two peptides with a MASCOT score > 95% were considered as significant.

Prokaryotic test

Females from six thelytokous strains were fed a solution of 1% tetracycline and 10% sucrose during four generation. One arrhenotokous strain and curing of mated females of the parasitic wasp *Nasonia vitripennis* infected with *Wolbachia* (LAB II strain) served as a control. In addition, we performed a PCR test for the presence of microorganisms in the ovaries of arrhenotokous and thelytokous females using prokaryotic and *Wolbachia* specific primers. Several organisms were used as a control to the experimental procedures for establishing presence and absence of *Wolbachia* and other prokaryotes (*Wolbachia*-infected and cured *Nasonia vitripennis*, *Wolbachia*-infected *Ephestia kuehniella, Escherichia coli* and *Lactococcus lactis V. canescens* DNA was obtained from ovaries of 5 wasps of each strain. Wasps were dissected under sterile conditions in PBS. DNA samples were used in PCR reactions with specific primers to detect the presence of *Wolbachia* (wspF 5'-TGGTCCAATAAGTGATGAA-GAAAC-3' and wspR 5'-AAAAATTAAACGCTACTCCA-3', named 81F/691R primers in Zhou *et al.*, 1998) or with 16S ribosomal DNA specific primers (B8F 5'-AGAGTTTGATCMTGGCTCAG-3' and U1406R 5'-ACGGGCGGTGTGTRC-3' according to Lane, 1991) to detect the presence of any prokaryotic endosymbionts. PCR reactions were performed in 1X PCR buffer magnesium free (Promega) with 2.5 mM $MgCl_2$, 0.2mM dNTPs (Roche), 0.2mM each primer, 0.4 units of Taq polymerase (Promega) and approximately 5ng of template DNA. The PCR profile was 1 cycle of 3 min at 94°C followed by 35 cycles of 1 min at 94°C, 30 sec at 55°C, 1 min at 72°C and a final cycle of 7 min at 72°C for wsp. For 16S amplification the PCR profile was similar except for the extension step in each cycle that was 4 min at 72°C.

Results

Transcriptome analysis

The cDNA digestion and specific amplification with 48 different AFLP primer combinations yielded a total of 2104 well resolved arrhenotokous and thelytokous fragments that ranged from 100 to 500 bp. Gels showed a good repeatability for each primer combination. The total number of differentially expressed fragments between both reproductive modes was 190 (9%) with an average of 3.96 different bands per primer combination. More fragments were found in the arrhenotokous sample (116 fragments) indicating that more transcripts are expressed in arrhenotokous than in thelytokous ovaries. On the basis of size, intensity and accessibility in the gel, 47 arrhenotokous and 49 thelytokous differentially expressed cDNA-AFLP fragments were selected and excised from the gel and reamplified. Subsequently, 92 fragments with good amplification products were subjected to

Table 6.1. cDNA-AFLP differentially expressed fragments in the ovaries of the two reproductive modes. Fragment size, Genbank accession number, reproductive mode in which the fragment was found and sequence identities and functions compared to other hymenoptera are listed. A= arrhenotokous, T=thelytokous.

cDNA-AFLP fragment	Size (bp)	GeneBank Accesion no	Reprod. mode	Sequence identities	Function	Species
VcE07	188	AY204698	A	Putative zinc-finger protein (FP)	Hormone response	*V. canescens*
VcA11	170	GD254135	A	Transducin-like enhancer protein 4	Co-repressor	*N. vitripennis*
VcC04	127	GD254143	A	LD23562p Uncharacterized conserved protein AY122162	Unknown	*N. vitripennis*
VcG06	158	GD254171	A	Transcription initiation factor IIA large subunit XP_001599853.1	DNA repair, regulation of transcription	*N. vitripennis*
VcG07	122	GD254172	A	Calreticulin, partial CG9429	Calcium ion binding	*N. vitripennis*
VcD08	194	DQ649290	T	Microsatellite Vcan 056	—	*V. canescens*
VcB09	153	GD254140	T	Adenosylhomocysteinase hydrolase CG11654	Coenzyme transport and metabolism	*A. mellifera*
VcE08	150	GD254159	T	Flocculin	Flocculation	*N. vitripennis*

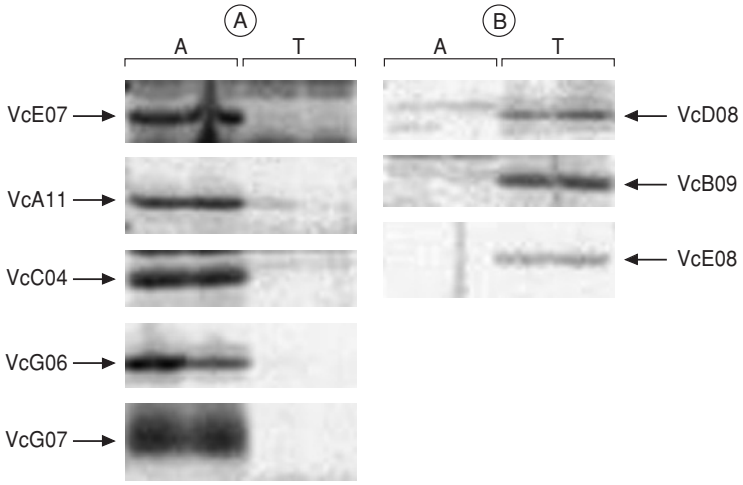

Figure 6.1. Enlarged view of polyacrylamide gels with differentially expressed cDNA-AFLP fragments. (A) Fragments present only in the ovarian tissue of arrhenotokous females. (B) Fragments present only in thelytokous females. The top row indicates the reproductive mode (A = arrhenotokous, T = thelytokous) loaded into each lane.

sequence analysis. Fifty-six fragments (61%) yielded reliable sequence data, although 26 fragments were excluded from further analysis because of their small size (<120 bp) they might represent artefacts.

Thirty cDNA sequences (17 arrhenotokous and 13 thelytokous, see Appendix I) were submitted at the National Center for Biotechnology Information (NCBI) site (http://blast.ncbi.nlm.nih.gov/Blast) in search for homologies. Twenty-eight sequences did not produce a significant result since their closest homologous gene had an E-value higher than $1x10^{-6}$. A direct nucleotide search revealed two V.canescens matches: fragment VcD08 to a microsatellite flanking region and fragment VcE07 to a zinc finger protein (Table 6.1). The small size of the other cDNA sequenced fragments (from 120 to 300 bp) prevented accurate sequence matches after a direct search for gene homologies in the NCBI database. We narrowed the search using NCBI program BlastX with the *V. canescens* a translated nucleotide sequences against the protein databases of *Apis mellifera* and *Nasonia vitripennis* which are the only two hymenopterans whose genomes have been sequenced. This search produced 6 significant hits. Sequence details and identities of the 8 differentially expressed transcripts are shown in Table 6.1 and their differential expression between the two reproductive modes is shown in Figure 6.1. These transcripts are putative orthologs of structural genes (Adenosylhomocysteinase, Flocculin), one chaperone (Calreticulin), two regulatory genes (Transducin-like enhancer protein, and transcription factors IIA) and one unknown protein.

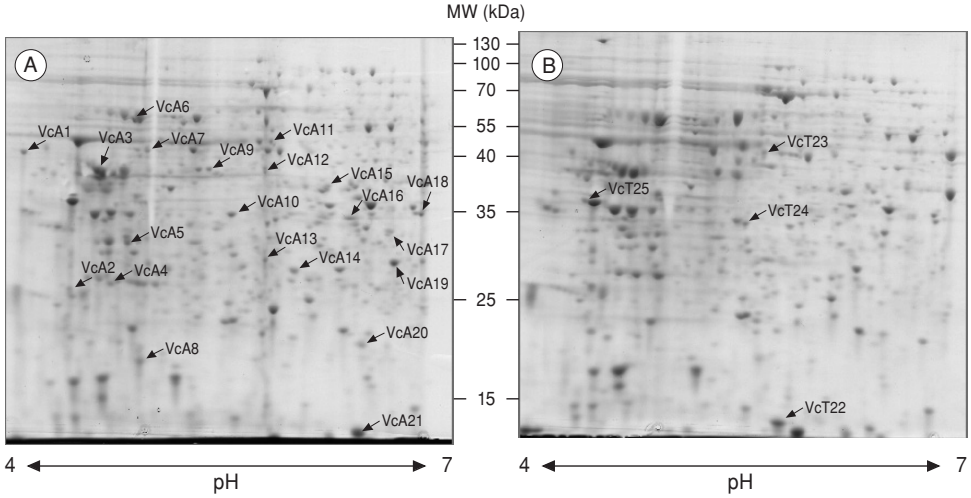

Figure 6.2. Comparative proteome analyses of arrhenotokous (A) and thelytokous (B) ovaries. Arrows indicate twenty-five proteins differentially expressed between the two reproductive modes that were separated on pH 4–7 IPG strips and identified by MS. A molecular weight (MW) marker with fragment sizes in KDaltons is indicated in the middle of the figure. Details are listed in Table 6.2.

Proteome analysis

Many ovarian proteins were visualized by two-dimensional gel electrophoresis at pH range 4-7 and differentially expressed spots between arrhenotokous and thelytokous ovaries could easily be identified. We chose this pH range because in a previous pilot study which included pH 3-6 we saw that the pH range 4-7 was more informative than the first one. Twenty one (84%) of the differentially expressed protein spots were only present in the arrhenotokous sample and 4 only in the thelytokous, in agreement with the observed more abundant RNA expression in arrhenotokous ovaries. The differentially expressed proteins are indicated with an arrow in Figure 6.2.

A local database for *V. canescens* was constructed comprising all 14 published *V. canescens* proteins in Genbank, the 30 RNA sequences from this study and protein sequences derived from the mass spectrometry analysis. Mass spectrometry results were compared to this database to check for specific matches. One of the protein spots was found to be the Virus Like Particle protein VLP2 (Reineke *et al.*, 2002).

The possible function of eight out of the 25 differentially expressed protein spots could be determined by comparison with the *N. vitripennis* and *A. mellifera* protein database from NCBI. The best matches for each spot are listed in Table 6.2. Two are homologues of structural proteins (tubulin, actin and the ribosomal protein S12), two have a regulatory function (RNA recognition motif) and one is a chaperone involved in posttranscriptional modifications.

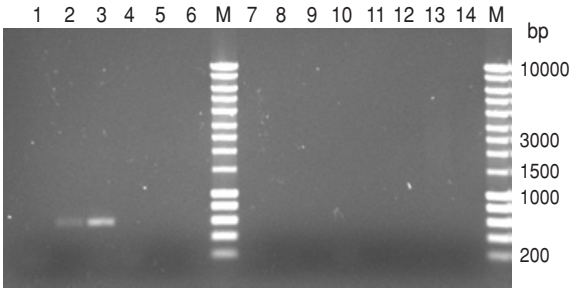

Figure 6.3. PCR amplification of a Wolbachia specific DNA fragment. 1. *L. lactis*; 2. *N. vitripennis* infected with *Wolbachia*; 3. *E. kuehniella* infected with *Wolbachia*; 4. *L. lactis-V. canescens* DNA mix; 5, 6, 7 arrhenotokous *V. canescens*; 8, 9, 10 thelytokous *V. canescens*; 11. *N. vitripennis* cured of *Wolbachia*; 12. Blank; 13. E. coli; 14. Blank; M. size marker.

To test for presence of endosymbionts that could manipulate the reproductive behaviour in thelytokous *V. canescens* females we carried out an antibiotic treatment with tetracycline during four generations (645 females were treated). As a result, no reversal to sexual reproduction occurred shown by the absence of males during the four generations. In both arrhenotokous and thelytokous *V. canescens* samples we observed no PCR products using neither *Wolbachia* specific primers (Figure 6.3) nor prokaryotic 16S primers (data not shown).

Discussion

Using a transcriptomics and proteomics approach we have identified several transcripts and proteins that show a presence-absence differential expression in the ovaries of the two reproductive modes of the parasitic wasp *Venturia canescens*. We compared RNA transcripts using 48 *TaqI/MseI* primer combinations which yielded 2104 ovarian cDNA fragments, from which 190 showed putative different expression in the two reproductive modes. Expression products are more abundant in the arrhenotokous samples for both mRNA transcripts and proteins which can be interpreted as arrhenotokous oogenesis requiring more ovarian gene products than thelytokous development. An arrhenotokous egg may require extra transcripts and proteins for processing the sperm nucleus after fertilization. Thelytokous females appear to have lost the function of egg fertilization possibly due to the action of Muller's Ratchet (Carson *et al.*, 1982; Pijls *et al.*, 1996; Arakaki *et al.*, 2000; Pannebakker *et al.*, 2005; but see Schneider *et al.*, 2003) even though they still mate and become inseminated (personal observation). Hence, differences in RNA and protein content between arrhenotokous and thelytokous females may not be causally related to diploidy restoration in thelytokous eggs but rather be due to

Table 6.2. Differentially expressed protein spots in the ovaries of the two reproductive modes. Genbank accession number, total ion score and protein identities. Spot number indicates reproductive mode: A = arrhenotokous, T = thelytokous.

Spot	Accession No.	Total ion score	Protein identities	Function	Species match
A.02	XP392313	208	β-Tubulin at 56D CG9277	Structural constituent of the cytoskeleton	*A. mellifera*
	gi|156538985|re	208	β-1 tubulin, partial	Structural constituent of the cytoskeleton	*N. vitripennis*
A.12	XP392899	58	Hsp60	Chaperon, posttranslational modification, protein turnover	*A. mellifera*
	gi|156541962|re	58	Chaperonine GroEL	Chaperon, posttranslational modification, protein turnover	*N. vitripennis*
A.13	gi|156551934|re	199	hypothetical protein	Unknown	*N. vitripennis*
A.18	XP392465	324	squid CG16901-PC	RNA recognition, RNA binding protein, gamete generation among other	*A. mellifera*
	gi|156554677|re	215	heterogeneous nuclear ribonucleoprotein	RNA recognition, RNA binding protein	*N. vitripennis*
A.19	XP623619	275	Actin-87E, 88F, 5C	Structural constituent of the cytoskeleton, cell motility	*A. mellifera*
	gi|156541054|re	275	RE12057p isoform 1 (Actin)	Structural constituent of the cytoskeleton, cell motility	*N. vitripennis*
A.20	XP001123058	67	Rbp1-like CG1987-PA	Nucleic acid binding, RNA splicing	*A. mellifera*
	XP397273	67	atypical protein kinase C CG10261-PA	Oocyte axis determination among other functions	*A. mellifera*
	gi|156552744|re	66	eukaryotic membrane protein	Unknown	*N. vitripennis*
A.21	gi|156546944|re	172	ribosomal protein S12	Structural constituent of the ribosome, translation	*N. vitripennis*
T.32	gi|156538110|re	56	alpha NAC	Protein binding, involved in oogenesis	*N. vitripennis*

other processes in which these two types of eggs differ, such as the ability of fertilization.

Results of the RNA and protein analyses show that the differences in maternal products are present as both mRNA and proteins. The identified transcripts do not match the differentially expressed proteins within a reproductive mode because the mRNAs have not yet been translated. It is known that many maternal mRNAs that code for proteins in the oocyte are present in the egg as ribonucleoprotein particles stored in the cytoplasm. These RNAs are expressed and translated later during early embryogenesis until zygotic transcription begins (Lovett & Goldstein, 1977).

We have identified a spot in the arrhenotokous sample as the previously described Virus Like Particle 2 of *V. canescens* (VLP2, Reineke *et al.*, 2002) which confirms the reliability of our study. VLPs are virus-like particles that cover the egg surface preventing the egg to be detected by the immune system of the host (Feddersen, *et al.*, 1986; Hellers *et al.*, 1996). The VLP-2 aminoacid sequence has high similarity to a RhoGAP domain (Reineke *et al.*, 2002) that activates small GTPases that control cellular morphogenesis through the reorganisation of the actin cytoskeleton (Bourne *et al.*, 1991). It has been suggested that RhoGTPases are also involved in animal cell mitosis in microtubule mediated processes such as spindle orientation and chromosome congression (Narumiya & Yasuda, 2006). Actin, which is related to spindle formation, is also differentially expressed in this study. In *Sacharomyces cervisiae* actin cables are known to act as rails for microtubule recruitment at spindle target sites (Gundersen & Bretscher, 2003). Hence VLP2 and actin are two possible candidates for causing the difference in chromosome segregation in arrhenotokous and thelytokous eggs.

A protein that shows expression in arrhenotokous but not in thelytokous ovaries is tubulin (*Tub56D*). Tubulin is the main component of the microtubules and plays a role in forming the mitotic spindle. In normal fertilized oocytes, sperm provides the centrosomes which together with the maternal pericentriolar material organize the microtubules which will move the oocyte pronucleus towards the centre of the cell (Morito *et al.*, 2005). In Hymenoptera, male arrhenotokous parthenogens develop from unfertilized eggs. These eggs do not receive paternal centrosomes, but are able to produce them *de novo* from maternal tubulin sequestered in the accessory nuclei (Ferree *et al.*, 2006). This is consistent with our observation of tubulin over expression in arrhenotokous ovaries and with the hypothesis that more tubulin is needed to process the paternal centrosomes if the egg is fertilised. On the other hand, thelytokous eggs do not accumulate tubulin and must have an alternative way to circumvent the lack of paternal centrosomes inheritance. A likely explanation is that unfertilised thelytokous *V. canescens* eggs can assemble the spindle from already existing cytoplasmic centrioles, as has been described for the two hymenopteran species *Muscidifurax uniraptor* (Riparbelli *et al.*, 1998) and *Nasonia vitripennis* (Tram & Sullivan, 2000).

Recently, genes related to thelytoky have been identified in aphids and in the honey bee. In the first one, using differential gene expression techniques Cortés *et al.*, (2008) identified, among others, a tubulin gene similar to *Tub56D* involved in the cascade leading to sexual phenotypes. In the honeybee, Lattorff *et al.* (2007) identified the transcription factor *grainy head* as most likely candidate for the *thelytoky* gene in *Apis m. capensis*. Similar to *A. m. capensis, V.canescens* achieves diploidisation of the thelytokous eggs by central fusion automictic parthenogenesis (Verma & Ruttner, 1983; Beukeboom & Pijnacker, 2000), but aphid eggs undergo apomictic parthenogenesis. Thus, aphids and the honeybee have two different cytological mechanisms of parthenogenesis which likely also have different genetic bases. Although the cytological mechanism in *V. canescens* resembles that of *A. m. capensis*, the absence of *grainy head* in our differential analysis suggests yet another genetic basis in *V. canescens*.

Of the other genes and proteins differentially expressed in our study, it is difficult to asses if and how they play a differential role between the two reproductive modes. Further study is needed to confirm whether the expression of these transcripts and proteins is functionally related to reproductive mode.

The results of Beukeboom and Pijnacker (2000) together with the antibiotic and PCR results presented in this study indicated that it is not likely that thelytoky is caused by endosymbionts in *V. canescens*. But studies like this one may present potential pitfalls (Weeks *et al.*, 2003) and one has to be cautious when interpreting negative data. But even if thelytoky would be induced in *V. canescens*, it would still be interesting to know what are the differences that cause diploidisation of thelytokous eggs and not of arrhenotokous ones.

To our knowledge this is the first study that compares the expression of the transcriptome and proteome in arrhenotokous and thelytokous oocyte development. We have identified several genes that are differentially regulated in arrhenotokous and thelytokous ovaries which may be causally related to diploidy restoration, fertilization or to other process that differ in arrhenotokous and thelytokous eggs. Further studies on the function of the candidate genes are needed to resolve this. Hopefully, our study will encourage more molecular studies to elucidate the genetic mechanism of parthenogenetic reproduction in insects.

Acknowledgements

We thank Henk Bolhuis for help with the prokaryotic amplification, Rasmus Larsen for supplying *Lactococcus lactis* DNA, Loes Drenth for help with the Bradford quantification of proteins and Carlos Bernstein for providing some of the wasp strains. This work has been made possible with support of the Netherlands Proteomic Center. The AFLP® technology is covered by patents and patent applications owned by Keygene N.V. AFLP is a registrated trademark of Keygene N.V.

Appendix I. Sequences of the cDNA-AFLP excised bands.

>VcA05
TACGAACACGAAAATGATCGCTGTCGTATGGTGCTTCCAGTTCTTCCATAACC-
GAATCCGGATCATCAACAAGTTCTTCATGTNTGTCTGAGAAGTGAATAATCGTTG-
TATTGTTCTCCGTGTAGTGTCATAGATGAACCCGGTTCATTAGCGACGGATCCTTC-
CTGCCCTGTGTTAA

>VcA06
TCGAACAACAACAACAACATTCTCACGCTCTATATACTTCCTACATCGTGACG-
TATTTCCGCTCACTCACACCCACACGGCAAATATCATACACACCGACATATTCT-
CATTTGGTAATAGTCATCACAACGACTTACGACTTATTGGGAAAATAGCTTTTG-
TACTTATTCATAACTAAAGACGAAGGTGTTAA

>VcA11
TCGACACTGGGACAACTCGTTCGTGCGTTGCATCGCGCGGATTACCGGACTTTAGC-
GAAACACGATAACGTGATCTGTCATTTCCAATTGTGAACGAGAACTACTCCCC-
TATAGCCACACGATAGAGGCCTATGGAAACTTTCTTCGTCTTCTCTTCTCCTGCAT-
GTTAA

>VcB02
TCGACGTTATCATAATGCTATTTTGGGCGTGAAAACAGTCCAGCATGCACTTC-
CAATCAGTTTGGATTTGACGATGAAGATGGTCCGTCGGGGACAGAACCGCCG-
GTTCAATATAGGAATGCCAGCGGCCGTCGGGGTAGCCGAAATTGGGTTAA

>VcB09
GCGTACGAACATGAAGATAGTATGAAATGGTGCGCACGC-
GATTCGTCGTTTTGATTTTCCGAAATTTCAACTAATATTGTAGT-
GATCGGGTTTGAAAGGTCCTTCGGCTGGTATGTCAAGATACTTGGCTTGGTCTG-
TAGTGAGTTTCGTTAA

>VcB10
ACTGGTACGAACGCGACTCTCTATGCTACAAATTGATCATCCCACGATTATTAGTG-
TATTTTTTTATACAATTCATATCTGCTACAGCCTTCCCCATTTTTATCGCTC-
CGTCGTCGCGCAAAATACGAAATCGCGATTATTTTTTTCGTATCGTTAA

>VcB12
ACGAACTTCAACAGCTTCGGAAGTATTTCTCACAACAGTTTCGTCTATAGATTGAAT-
TACAGCATTCTCAGTAGCATCGTAAAACCGCGCGATACCAACTGCGCCTTTCTTAG-
CAAAAGCACAAAGGTCACCACACACATGGGTGATCTTCTTTTCTTTTTTCGGCT-
TACTGATACTGTCAAACACGCGCGTTAA

>VcC04
TCGAACATGGAACGAAATCTCTTCTCTATTCTTACCTTGCGCTCGCTTTTCGGGATC-
CTTCGGAATTGTATGGCCTTGGCGATATGCGCGACCCGCTTGTCTTTCAAC-
CTTTCGTCCATGACGTTAA

>VcC07
TCGAACAGGGAATCTGATACGTCATGATACCAGACTTGAGATCTATTCAATAC-
TATTTTCGTGATTCAATACTAGTTATGCGTAGGTATGTTTTTC-
CTTTCTTTCTTTCTTTCTTTCTCTCTATCCTTCTTTCTGGTAGTTAA

>VcC09
TCGAACAAATGGAAATGACTGTATAAAATCAGGTGTAACCGTGATACGTGACATA-
GAAAAAAAATGAAGTCCTAATACATCAACAACACAGAATTATGTTT-
TATCTTTTTTTTCATTTTGTACCAATATTATACATTATTCCGTAATAATTACGACGAA-
CAAACAGAGCACGAGTTAA

>VcD01
TCGAAGGAAAGAAATCATTCTCTGAAGTACGTGGCGCGTATAAGAATGTG-
GAAAATGATATAAACATGTTTTTATCTGACTGTGGGAATTGTGAATTTTTGT-
TAATCCAGCGCCACATAATCCGTGTTAA

>VcD03
TCGAGGAGAACGAATACATCCATGGCACATTGCCGGGACAACACTTTGTAGTATT-
TATATTGGTCTCATATAGATATATGTATAATGTTTATGCTTGGATAAATCGGCACTCCT-
GTTAA

>VcD08
TCGAGNCGTCCATGCTAACAAACAATTGCTTTCGTTGATTTTCGGAAAAGTAAT-
CAAACTACATAATTACCAAAAACCGCGCGTCAAATAGTTATTTTTTTCCTTCAATTT-
TATATTTTTTATATATATATATATTTTGTTTTTATATCTACCTTGAAACG-
GTAATAAACTCGGTTACCGAAACTGTTGGTTAA

>VcD11
TCGACGTTTCATAATGCTATTTTGGGCGTGAAAACAGTCCAGCAGGGCACTTC-
CAATCAGTTTGGATTTGACGATGAAGATGGTCCGTCGGGGACAGAACCGCCG-
GTTCAATATAGGAATGCCAGCGGCCGTCGGGGTAGCCGAAATTGGGTTAA

>VcE01
TCGAGTGACCCGCGATGAACGAGTAACGAGATGCTTTCAGATAATTGCGAAAAT-
CAAGAATTTCCAGTGTAATGTACTCTCAAGGCTAGATTTACGACAGAAATATACT-
GACAATTTCAGAATAACTGCAACTTTCGGTTAA

>VcE03
TCGAGATGAAATCTGATCCGGGCGATATACTTGACGCCGGCGTACTCAACGGTCT-
TATCACTATTATCATGAAAATTGTGATTCCGGAGTGTCGCTGGTTGTTGCT-
GATGGGATGAATTGTGGCGATGAGGGACCTTTTGATGCAGAAGAAATGTATAG-
GTTAA
>VcE04
TCGAGGTCTTTCAGGATTCGTACTCAAACGATTTCAAAAAGCAGAACAATAATT-
TACTCTCGTCTTGACACTCTTTTCGTCTGCTGTATGTGCAACGAAATGAAACTAGCG-
TATTGCGAAGGTGTAAGCTCTGTGTCTCCAGTAGTGACTAATTCGTTAA

>VcE05
TTTATAAACACGGTTTGCGACGTTCAGTTTCGTTGAAGATAAGACAATTTTTATTCT-
GCAACGAACGATGTGATTCAGTGCCTATTACGGAAACAATCTC-
TATCTCTTTTCATATTTTTCCCTTTACTCTCCCTCTCTCTCTCTCTCTCTC-
TATCTCTTTCTTTCCTGATTCGTTAA

>VcE07
TCTGAAGGGTACCTGNCGAGTTCTGAAGCCGGCAATAGAAATCAC-
GAACGTTCTTTGTCGCGGTGAATGCAGATCGGTCAGCCAGCACGAGA-
CAACAGGTTCTGAAGAAAATTTTACGCCACAGAAGCGAAAACAATGTCAACC-
GATTCACCGTGGTCGGGTTGATCGCTGCTGCGATGCGGCGTTAA

>VcE08
CATCAAAGTATCATTGCGTACTGGCAAAACTGAATTTCGTACGCGATTAGTGAGC-
GATTGCATTACGATCATTTGATTCTGCTGTGCGTTACCCAAAATTATTGGATTTGAC-
GATTGAGCTTGGGCACTCACGAGTATATTTGGCGTTAA

>VcE11
AATGGTTTATGTAAAATAGAGTTGTTGTTGTTTCATCACCGTTTTGCACAAA-
CACTGGCTGATTGTAACCAATGAATTTCACAACGTGATCTTCCTCTATAAT-
CAATTCTGGGCGACGTTAA

>VcF04
TCGAATGCCAGTTTATGTTTCTCATCAACGAGAAACTGTACTTTTGTCCAAAAG-
CAAGAGTTACGCGAAGCCTCAAGAGGGCTCGTAAAGACCG-
CAAACTCTTGAATCGCCCCACTGAAGTGGAGGCTTTTTCCCCTCTATGTTATAGA-
GAGGAATTTGAGGTTGAACCAGTTACT

>VcF12
TCGAAACTGAAGAGTGCCCGAGCCCCTNTCTCTGGAATGAAGAAAGCAGCGCC-
GACGTAGTCTGCATTTTTTGTTTTGGATCCGTCTGTGAAGATCAGAGTCCAATTGT-
TACTACTATGGTGCCTGTATCATGTTAA

>VcG05
ACGACAACAGCTAGTTGAGTGGTTCAGTGCTGCGCGATACAGTTCG-
TAAAAATAATTCATGAGAATGGCTCTGCACAGATTGAAAAAATATTGAAAATCTA-
CAAGTCAGTGGCACAGAATCCTGCTTTGATATCTGCCAGAATCACTTGGGTTAA

>VcG06
ACGACAAGATTACAGAAGCAGAAACAAATGGAAATTCTACTTGAAGGATGGAAT-
TATGAACTTGAGCGGCAAGGATTACGTTTTCCAGAAGGCAAATGGTGACGCG-
GAATGGTGAATTGCGGAAAAAATCTTTTTTCGTGATATTGCCACGGGGTTAA

>VcG07
TCGAAACCTGACTACAGGGACGTGGGTTCATCCTGAAATTGACAATCCCGATTA-
CACTCCCGATACCGAACTTTACAACGAGATGAAATTTGTGCTATTGGTTTTGAC-
CTCTGGCAGGTTAA

>VcG08
TCGAAGCTTCATCGGTCTTGCTAAAGAGTCAAATCTGCGAAGAT-
TATTTGAAAAACAGAACGTTTCAACAATCTATCGCAACGTTTTTATTG-
GTTTGTGGGAAAAAGAAATTATATGGGAAGAAATTACTGTAACGATTAGTCATTAG-
CATAAATTTGCTTGAAAAAAAATCAATGAACAATCGGAAAGCGCTTTTTCGTTAA

>VcG11
TTCCATACGGTGAACACGGAGCACTATACCGGAAATAAGATAAATCTAAATACC-
CTTTTGAATTTCAATCACTCTCGACTTGAATAAATGAATCAGC-
TACTTTCAATATTTCGTTGCGTTAA

>VcG12
TCGAATAGAAATCATAGAGGGAGAAGAGAGGGCCCGCTCGGAGGTCCA-
GATCAGTGCGAATGCGTTCACTTATTATTGTGTTTCAAGTAAAATTGCA-
GATAGTTTGTGGAACGAGGATGAGTGCGGGGGAGCGAAAAGTGAAAGAAGAG-
GATCGCAAATTCTAGCAGGCACTTAGTAATGAGATAAGATGGCAGTTGCGCTAGT-
TAGACTTGTCCAAAAATGGCGT

>VcH10
GTACGACAGCATTTTGATTTCCAATTCATAAGCGGAGGTTGAATTTCT-
GTTTCTTTCTGGCGAATATAAAAAAAGAAATTGACGCTACGTCTTGCGATAAT-
CACTGGATATTCCTTTGAGGAATTCCGTAAACAAACGGTCCTTTTCCAGTTTTTCTG-
GAGTTAA

>VcH12
TCGAGTGCACTATGGAATGTGAAATTCTACGACAGTCATCCAATAATCATGACTGT-
GATTTTGAAACGAACAAAAATCCGATCGTACGACGAAGGATTCGAAAGT-
GATTCGCGCACTTTTTCAAGTTCAGTATCCATCACTTTACATTCAGTTGAATTCAAA-
GAGGCCTTGGCTCTTCAAAACTTTACATTGTTATCATGTTTCTGAAACCAAGTGCT-
TACCTGAGGGTTAA

Appendix II-A. MS results for the *A. mellifera* comparison. For each spot number and reproductive mode (A = arrhenotokous, T = thelytokous), Accession no. of the matched *A. mellifera* contig are indicated as well as protein and peptide information extracted from the MS analysis.

		Protein information					Peptide information				
Spot	GenBank accession No.	MW	PI	Pep. Count	Score C. I. %	Total Ion Score	Start Seq.	End Seq.	Sequence	Ion Score	C. I. %
A.01	Amel|GB17896-PA	96833.1	9.8	14	86.07						
A.02	Amel|GB10275-PA	50143.1	4.8	8	100	208	47	62	INVYYNEASGGKYVPR	73	100
							63	77	AILVDLEPGTMDSVR	44	99.982
							78	103	SGPFGQIFRPDNFVFGQ	12	76.287
							104	121	GHYTEGAELVDSVLDVV	26	98.984
A.03	Amel|GB10996-PA	36297.1	9.6	9	0						
A.04	Amel|GB11957-PA	39589.1	6.5	6	0						
A.05	Amel|GB16120-PA	68811.7	9.1	13	87.30						
A.06	Amel|GB12151-PA	116164.2	8.9	20	99.22						
A.07	Amel|GB10836-PA	91146.3	5.8	5	99.92	61	587	597	NALEEYVYDLR	61	100
A.08	Amel|GB19168-PA	21229.3	9	6	23.47						
A.09	Amel|GB17681-PA	41778.8	5.3	8	100	64	240	255	SYELPDGQVITIGNER	64	100
A.10	Amel|GB17457-PA	13409.5	5.2	6	74.07						
A.11	Amel|GB18292-PA	45734.2	8.6	14	97.64						
A.12	Amel|GB18969-PA	60374.9	5.6	8	100	58	56	67	NVILEQSWGSPK	24	98.407
							92	116	LVQDVANNTNEEAGDGT	34	99.833
A.13	Amel|GB19952-PA	53506.3	8.7	12	62.52						
A.14	Amel|GB12732-PA	27937.9	5.8	2	96.66	49	21	29	VFQVEYAQK	49	99.995
A.15	Amel|GB12453-PA	41769.6	5.3	6	83.64	25	30	40	AVFPSIVGRPR	4	0
							240	255	SYELPDGQVITIGNER	21	96.791
A.16	Amel|GB13436-PA	28125.9	10	6	0						

Appendix II-A. Continued

Spot	GenBank accession No.	Protein information MW	PI	Pep. Count	Score C. I. %	Total Ion Score	Peptide information Start Seq.	End Seq.	Sequence	Ion Score	C. I. %
A.17	Amel|GB18857-PA	28356	9.9	9	100	52	11	21	IYVGNLPPDIR	17	94.118
							45	66	GPPFAFVEFDDPRDAEDAVHAR	10	68.628
							124	143	VLVTGLPPSGSWQDLKDHMR	25	99.087
A.18	Amel|GB15796-PA	31916	5	6	100	324	55	71	KLFVGGLSWETTDKELR	62	100
							72	86	DHFGTYGDIESINVK	109	100
							94	104	SRGFAFIVFAK	35	99.853
							96	104	GFAFIVFAK	67	100
A.19	Amel|GB18527-PA	41757.8	5.3	7	100	275	30	40	AVFPSIVGRPR	23	98.795
							86	96	IWHHTFYNELR	77	100
							240	255	SYELPDGQVITIGNER	92	100
							293	313	DLYANNVLSGGTTMYPGIADR	82	100
A.20	Amel|GB13703-PA	11794.8	10	3	100	67	43	48	NVWVAR	19	94.269
							49	69	NPPGFAFVEFEDPRDAEDAVR	40	99.954
A.20	Amel|GB19525-PA	75063.8	8	4	99.96	67	39	44	NVWVAR	19	94.269
							45	65	NPPGFAFVEFEDPRDAEDAVR	40	99.954
A.21	Amel|GB15605-PA	15263.6	5.5	2	100	106					
T.31	Amel|GB19016-PA	416068.3	4.8	52	99.33		232	243	AVTLMSADLSKR	1	0
T.32	Amel|GB18087-PA	22669.6	4.7	3	99.81	56	97	111	SKNILFVINKPDVLK	30	99.716
T.38	Amel|GB17447-PA	21751	8.2	3	0	21					
T.39	Amel|GB14067-PA	158927.4	9	33	100	10	1130	1142	ESSPMLDAKNDIR	8	58.05
							1230	1251	IWGTDSNMSSLQSLEGSDVEGR	10	74.19

Appendix II-B. MS results for the N. vitripennis comparison. For each spot number and reproductive mode (A = arrhenotokous, T = thelytokous), Accession no. of the matched N. vitripennis protein are indicated as well as protein and peptide information extracted from the MS analysis.

Spot	GenBank accession No.	Protein information			Peptide information	
		MW	Pep. Count	Total Ion Score	Sequence	Ion Score
A.01	gi|156543439|re	138838	1	12	QEILCLNVFSR	12
A.02	gi|156538985|re	37282	7	208	INVYYNEASGGK	38
					AILVDLEPGTMDSVR	44
					AILVDLEPGTMDSVR	21
					INVYYNEASGGKYVPR	73
					GHYTEGAELVDSVLDVVR	26
					GHYTEGAELVDSVLDVVRK	15
					SGPFGQIFRPDNFVFGQSGA GNNWAK	12
A.03	A5I9E8|A5I9E8_N	235544	1	20	YFVSPTLLR	20
A.04	gi|156540582|re	66952	1	16	LNDAWPLLGEKLNR	16
A.05	gi|156553204|re	51520	1	17	VKNQFLMPDSVK	17
A.06	gi|156537462|re	60384	1	25	LYQLKALQDASR	25
A.07	gi|156548968|re	92567	1	61	NALEEYVYDLR	61
A.08	gi|156549887|re	16911	4	68	TYGEFPETVR	36
					NLKTYGEFPETVR	10
					VWINQGDIILIGLR	10
					ELVFKEDGQEYAQVTK	11
A.09	gi|156541054|re	41758	1	64	SYELPDGQVITIGNER	64
A.10	gi|156552051|re	104204	1	17	IGKMHNYFSVESAR	17
A.11	gi|156546835|re	39729	1	22	MPPQIVAIDQDLNR + Oxidation (M)	22
A.12	gi|156541962|re	60549	2	58	NVILEQSWGSPK	24
					LVQDVANNTNEEAGDGTTT ATVLAR	34
A.13	gi|156551934|re	36393	4	199	VNQIGSVTESINAHK	49
					YNQILRIEEELGAAAK	40
					VNQIGSVTESINAHKLAK	15
					SGETEDTFIADLVVGLSTGQIK	96

Appendix II-B. Continued

Spot	GenBank accession No.	Protein information			Peptide information	
		MW	Pep. Count	Total Ion Score	Sequence	Ion Score
A.14	Gi|156555588|re	84422	1	20	QVANLEESLGR	20
A.15	gi|156541054|re	41758	2	25	AVFPSIVGRPR	4
A.16	gi|156540998|re	204255	2	27	VESPIATYIKQGAIPK	17
A.17	Q95VR1|Q95VR1_9	40049	2	142	YGNEEIDQDDPVYENVNLR	57
					SQTVNQGQPELLYASLDLP VTNR	85
gi|156545752|re28739		3	52		IYVGNLPPDIR	17
					VLVTGLPPSGSWQDLKDHMR	25
					GPPFAFVEFDDPRDAEDAV HAR	10
A.18	gi|156554677|re	35841	4	215	GFAFIVFAK	67
					SRGFAFIVFAK	35
					LFVGGLSWETTDKELR	52
					KLFVGGLSWETTDKELR	62
A.19	gi|156541054|re	41758	4	275	AVFPSIVGRPR	23
					IWHHTFYNELR	77
					SYELPDGQVITIGNER	92
					DLYANNVLSGGTTMYPGIADR	82
A.20	gi|156552744|re	56841	3	66	NVWVAR	19
					NPPGFAFVEFEDPR	9
					NPPGFAFVEFEDPRDAEDAVR	40
A.21	gi|156546944|re	15316	3	172	NALIHDGVVHGLHEAAK	106
					DFGEETPAKDVLMEYLK	66
					DFGEETPAKDVLMEYLK	13
T.31	gi|156550807|re	101734	1	21	FMLRDVIELR	21
T.32	gi|156538110|re	22389	2	56	NILFVINKPDVLK	26
					SKNILFVINKPDVLK	30
T.38	gi|156546028|re	45497	1	31	KTKSIERNR	31
T.39	gi|156550225|re	57869	1	29	KKALELSAVYR	29