**Supplementary Discussion**

**Insights from non-toxin genes and pathways enhanced in the VG**

Here we provide a more detailed discussion about genes and pathways identified as differentially expressed in the VG, with some insights into the VG biology.

The gene enrichment analysis suggested an important participation of several peptidases and modifying enzymes that are likely to be important for the production of venom toxins. For example, X-prolyl aminopeptidase is a proline-specific metalloaminopeptidase responsible for removing N-terminal amino acids found adjacent to proline residues. Because of its specificity toward proline, it has been suggested that X-prolyl aminopeptidase is important in the maturation and degradation of peptide hormones and neuropeptides, including bradykinin (Cottrell et al, 2000). Another gene identified, glutaminyl-peptide cyclotransferase, catalyzes N-terminal pyroglutamate formation. This is one of the few enzymes involved in toxin biosynthesis that has previously been investigated (Pawlak and Kini, 2006) and it has recently been identified as secreted in venom (Wang et al., 2014), although it is unlikely to have toxin-related activity. Rather, since N-terminal pGlu is very common in many snake toxins (e.g. BPPs, 3FTx, metalloproteinases) glutaminyl-peptide cyclotransferase presumably confers a peptide protection against degradation of these toxins in the preys.

Chaperones are also highly represented in the VG. The two most expressed non-toxin unigenes in the VG correspond to calreticulin and 78 kda glucose-regulated protein (HSP70). Both proteins promote protein folding, help
in the assembly of oligomeric multiunit proteins and contribute to the identification of misfolded proteins. Relating to this, e3 ubiquitin-protein ligase, a member of the endoplasmic reticulum quality control (ERQC) system, was also among the highly VG expressed unigenes.

Protein disulfide isomerases (PDI) and other proteins involved in disulfide bond formation were found to be upregulated in the VG when compared to the other tissues sampled (Table S3) and even when compared with the pancreas (data not shown). The high expression of PDI in the VG was previously reported, but until now there was no information about its expression in other tissues (Junqueira de Azevedo and Ho, 2002, Zhang et al., 2006). PDI is a redox protein with independent chaperone activity and is responsible for correct disulfide bond assembly in the endoplasmatic reticulum (Freedman et al., 1994). Other functions of this protein, related to its ability to bind many targets by cysteine residues, have been reported, including biotechnological uses as a refoldase for recombinant proteins or snake PLA₂ (Yao et al., 1997). The S-S bridge arrangement is a striking feature of toxins, and is seemingly fundamental to the structure, stability and activity of almost all classes of snake venom toxins. Considering toxins are secreted proteins that have to survive many unfavorable conditions, from the high density storage in the venom gland lumen to distinct physiological environments in prey, disulfide bonds are of utmost importance to their stability. In addition to the contribution of PDI to this process, other proteins involved in disulfide bond assembly and the regulation of the redox state in the cell, such as selenoprotein M and glutathione peroxidase, were also found to be present in our set of differentially VG expressed unigenes.
Although few individual unigenes identified in our VG differentially expressed dataset (Table S3) were associated with glycosylation pathways, our gene enrichment test (Figure 4) highlighted the apparent importance of this pathway for venom biosynthesis. In particularly, we found that several subunits of the dolichyl-diphosphooligosaccharide-protein glycosyltransferase, which acts as part of an essential N-oligosaccharyl transferase complex that promotes N-glycosylation, contributed to this result. Glycosylation is a key feature of several toxins and it is at least partially responsible for the huge diversity of proteoforms observed in most snake venom proteomes (Zelanis et al., 2012). For example, the same toxin type may exhibit different levels of sugar addition, as we recently observed for a B. jararaca serine protease (Yamashiro et al., 2014, in press).

Vesicular traffic was also detected as one of the most enriched processes for the VG differentially expressed unigenes. This is perhaps not surprising, considering the VG is a secretory tissue whose epithelial cells are filled with secretory granules during its active state (Kochva, 1987). Putative genes that we identified involved in this function are sec61 and coatomer subunits. It is interesting to note that the exocytosis-regulating protein syncollin, which we found highly expressed in the B. jararaca pancreas transcriptome, was not detected in the VG. This implies that the upper steps of the secretory pathway in the VG may follow a different type of regulation to that of the pancreas. However, calglandulin, a protein that we previously proposed to be a regulator of vesicular traffic in the VG (Junqueira-de-Azevedo, et al., 2003), was only represented by a single read in the B. jararaca VG transcriptome, despite it
being shown to be highly expressed in the VG of several other snakes (St Pierre et al., 2005; Zhang et al., 2006).

Among particular unigenes highly differentially expressed in VG, some were unexpected. Lysophosphatidylcholine acyltransferase 2, besides being involved in the generation of 1-alkyl-phosphatidylcholine (PC), a major component of cell membranes, is also involved in the conversion of lyso-PAF into platelet-activating factor (PAF). The PAF can trigger inflammatory and thrombotic cascades and mediate molecular and cellular interactions between inflammation and thrombosis. (Zimmerman et al., 2002). Inflammation and coagulation cascades are critical features of B. jararaca envenoming, raising the possibility that the lysophosphatidylcholine acyltransferase 2 or, most reasonably, PAF could be secreted to venom. This possibility remains to be tested. Anexin 5 also has a high expression level in VG. This is one of the few extracellular annexins normally abundant in the ductal perfused organs. This protein is secreted in a non classical way and forms a shield on cell-surface phospholipids that acts as an indirect inhibitor of the thromboplastin-specific complex, which is involved in the blood coagulation cascade (Gerke et al., 2005; Bogdanova et al., 2012). Considering the high abundance of procoagulant elements in the venom, a protective role of this protein on the VG cell surface against these elements could be imagined.

New VG exclusive unigenes could be observed on Table S4. Dipeptidase 2, for instance, is a GPI anchored protein that digests a variety of free dipeptides (Habib et al., 2003) that has never been observed in VGs transcriptomes. It is not related to the venom dipeptidylpeptidase IV, observed among toxins, that acts by selectively removing the N-terminal dipeptides from
peptides with proline or alanine in the penultimate position (Ogawa et al., 2006). Besides two gamma PLA$_2$ inhibitor unigenes exclusively (and highly) expressed in the liver transcriptome, we found a third gene exclusively expressed in VG. These inhibitors are known to be present in snake plasma where they act as one of the key elements protecting the animal from envenoming by other snakes by efficiently inhibiting venom PLA$_2$ (Lizano et al., 2003). However, the occurrence of this kind of inhibitor in the VG is completely novel. Notably, the VG unigene we detected here shows considerable sequence differences to the $B$. jararaca liver forms and those previously known from blood plasma (Figure S13A). Phylogenetic analysis of these inhibitors revealed that the liver expressed genes group into two clades with all other snake sequences, indicating a fairly ancient duplication event (Figure S13B). Notably, the venom gland expressed gene shows no similarity to these typical forms and is found near the base of the tree.

Non-venom gland expressed genes that group phylogenetically with toxins

We found two other examples of non-venom gland expressed $B$. jararaca genes forming monophyletic groups with previously described venom toxins. The 3FTX gene BJARALL03759 was found to be selectively expressed in the pancreas (Figure 7A and Figure S11), yet it grouped with a 3FTX isolated from the venom gland of the homalopsid snake $Enhydris$ polylepis (Bayesian pp = 0.86) within the monophyletic snake venom 3FTX clade. However, when excluding this sequence from the toxin clade, Bayes factors analyses revealed only moderate support for its placement (Table S5), and therefore we cannot
exclude the possibility that this gene lies outside of the venom toxin clade and represents the physiologically expressed ortholog to venom 3FTXs.

The cobra venom factor gene BJARFIG00306 is selectively expressed in the liver of *B. jararaca* and was found to form a robustly supported monophyletic group (Bayesian pp = 0.99) with sequences obtained from the venom/venom gland of the elapid snakes *Ophiophagus hannah* and *Naja kaouthia* (Figure S2 and Table S5). Our character analyses suggest that this toxin type may have been recruited into the venom of snakes on two independent occasions, with our *B. jararaca* gene therefore representing a physiologically expressed ortholog to one of these groups. However, support for this hypothesis (Bayesian pp = 0.86) did not reach the significance threshold of the test and was only moderately higher than that supporting the alternative hypothesis of a single venom recruitment event, followed by a change to physiological expression in *B. jararaca* (Bayesian pp = 0.76) (Figure S2).

**References**

Bogdanova N, Baleva M, Kremensky I, Markoff A. 2012. The annexin A5 protective shield model revisited: inherited carriage of the M2/ANXA5 haplotype in placenta as a predisposing factor for the development of obstetric antiphospholipid antibodies. *Lupus* 21(7):796-798.

Cottrell GS, Hooper NM, Turner AJ. 2000. Cloning, expression, and characterization of human cytosolic aminopeptidase P: a single manganese(II)-dependent enzyme *Biochemistry* 39:15121-15128.
Freedman RB, Hirst TR, Tuite MF. 1994. Protein disulfide isomerase: Building bridges in protein folding. *Trends. Biochem. Sci.* 19:331–336.

Gerke V, Creutz CE, Moss SE. 2005. Annexins: linking Ca2+ signalling to membrane dynamics. *Nat. Rev. Mol. Cell Biol.* 6(6):449-61.

Habib GM, Shi ZZ, Cuevas AA, Lieberman MW. 2003. Identification of two additional members of the membrane-bound dipeptidase family. *FASEB J.* 17(10):1313-1315.

Junqueira-de-Azevedo ILM, Ho PL. 2002. A survey of gene expression and diversity in the venom glands of the pitviver *Bothrops insularis* through the generation of expressed sequence tags (ESTs). *Gene* 299:279-291.

Junqueira-de-Azevedo Ide L, Pertinhez T, Spisni A, Carreño FR, Farah CS, Ho PL. 2003. Cloning and expression of calglandulin, a new EF-hand protein from the venom glands of *Bothrops insularis* snake in E. coli. *Biochim Biophys Acta* 1648:90-98.

Kochva E. 1987. The origin and evolution of the venom apparatus. *Toxicon* 25(1):65-106.

Lizano S, Domont G, Perales J. 2003. Natural phospholipase A(2) myotoxin inhibitor proteins from snakes, mammals and plants. *Toxicon* 42(8):963-977.

Ogawa Y, Mamura Y, Murayama N, Yanoshita R. 2006. Characterization and cDNA cloning of dipeptidyl peptidase IV from the venom of *Gloydius blomhoffi brevicaudus*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 145(1):35-42.
Pawlak J, Kini MR. 2006. Snake venom glutaminyl cyclase. *Toxicon* 48(3):278-286.

St Pierre L, Woods R, Earl S, Masci PP, Lavin MF. 2005. Identification and analysis of venom gland-specific genes from the coastal taipan (Oxyuranus scutellatus) and related species. *Cell Mol. Life Sci.* 62(22):2679-2693.

Wang YM, Huang KF, Tsai IH. 2014. Snake venom glutaminyl cyclases: Purification, cloning, kinetic study, recombinant expression, and comparison with the human enzyme. *Toxicon* 86:40-50.

Yamashiro ET, Oliveira AK, Kitano ES, Menezes MC, Junqueira-de-Azevedo IL, Paes Leme AF, Serrano SM. 2014. Proteoforms of the platelet-aggregating enzyme PA-BJ, a serine proteinase from Bothrops jararaca venom. *Biochim Biophys Acta.* In press (doi: 10.1016/j.bbapap.2014.09.012).

Yao Y, Zhou Y, Wang C. 1997. Both the isomerase and chaperone activities of protein disulfide isomerase are required for the reactivation of reduced and denatured acidic phospholipase A2. *EMBO J.* 16(3):651-658.

Zelanis A, Serrano SM, Reinhold VN. 2012. N-glycome profiling of *Bothrops jararaca* newborn and adult venoms. *J. Proteomics* 75:774-782.

Zhang B, Liu Q, Yin W, Zhang X, Huang Y, Luo Y, Qiu P, Su X, Yu J, Hu S, Yan G. 2006. Transcriptome analysis of *Deinagkistrodon acutus* venomous gland focusing on cellular structure and functional aspects using expressed sequence tags. *BMC Genomics* 7:152.

Zimmerman GA, McIntyre TM, Prescott SM, Stafforini DM. 2002. The platelet-activating factor signaling system and its regulators in syndromes of inflammation and thrombosis. *Crit. Care Med.* 30(5 Suppl):S294-301.