Title page

Letter to the Editor

A plea for standardized nomenclature of snake venom toxins

John F Mulley[1] and Adam D Hargreaves[1]

1. School of Biological Sciences, Bangor University, Brambell Building, Deiniol Road, Bangor, LL57 2UW

John Mulley – j.mulley@bangor.ac.uk

Adam Hargreaves – adamdhargreaves@gmail.com

"…they have all one language; and this they begin to do: and now nothing will be restrained from them, which they have imagined to do. Go to, let us go down, and there confound their language, that they may not understand one another's speech."

Genesis, Chapter 11, verses 6-7

In the book of Genesis a spiteful God replaces a single, unifying language with a diversity of different ones, removing at a stroke mankind's ability to work together towards a common goal. We believe that the Tower of Babel myth has clear relevance to the snake venom literature, particularly to the nomenclature of snake venom toxins.

An established nomenclatural system for human genes, administered by the Human Genome Organisation (HUGO) Gene Nomenclature Committee (HGNC), has existed since 1979 (Shows et al., 1979) and similar systems also exist for mice (administered by the International Committee on Standardized Genetic Nomenclature for Mice) and zebrafish (administered by the Zebrafish Nomenclature Committee (ZNC)), as well as various other animal models. As a result, orthology, paralogy and the evolutionary relationships of these genes are clear for all to see. A nomenclatural committee was established by the International Society on Toxinology (IST) in 1992 after several exchanges in the 'Letters to the Editor' section of Toxicon in the early 1990's (Aird, 1990; Kaiser, 1990; Kumar et al., 1990), and more recently there have been suggestions for rational nomenclatural systems in spiders (King et al., 2008), scorpions (Tytgat et al., 1999), sea anemones (Oliveira et al., 2012) and centipedes (Undheim et al., 2014). However, despite over 20 years of asking, a nomenclatural system for reptile toxins has still not been established. The lack of widely-accepted, unifying standards for the nomenclature of snake venom toxins has resulted in huge diversity and disparity in the literature. In most cases the names assigned provide no insight into the evolutionary origins or relationships of the toxins, complicating comparative studies of venom composition, function or evolution both within and between species. For example, the similar sounding *ophanin* (accession AY181984, (Yamazaki et al.,

2003)), *opharin* (accession AY299475, Direct submission) and *ohanin* (accession DQ103590, (Pung et al., 2005)) from the king cobra (*Ophiophagus hannah*) are members of two different gene families – *ophanin* and *opharin* are cysteine-rich secretory proteins (CRISPs) and *ohanin* is a vespryn-like gene. Conversely, the differently sounding *pseutarin C* from *Pseudonaja textilis* (accession AY260939, (Rao et al., 2004)) and *trocarin D* from *Tropidechis carinatus* (accession DQ017707, (Reza et al., 2007)) are both *coagulation factor X* genes and *barietin* from *Bitis arietans* (accession FJ554635, (Yamazaki et al., 2009)), *apiscin* from *Agkistrodon piscivorus* (accession FJ554636, (Yamazaki et al., 2009)) and *cratrin* from *Crotalus atrox* (accession FJ554637, (Yamazaki et al., 2009)) are all *vascular endothelial growth factor F* (*vegf-f*) genes. The snake venom literature is replete with further examples, all of which obscure the relationships and functions of the toxins themselves – who would imagine that "cobra venom factor" which has been known since the late 19th century (Vogel, 1991) is part of the complement system from its name alone? Indeed, we have recently shown (Hargreaves et al., 2014a) that the *complement c3* gene has in fact been duplicated in cobras, giving rise to *cobra venom factor*, which should more accurately be called *complement c3b* to reflect its true evolutionary history.

We have previously suggested (Hargreaves et al., 2014b) that the "evolutionary characterization code" proposed by the Anolis Gene Nomenclature Committee (AGNC) for lizards of that genus (Kusumi et al., 2011) should be applied to snake venom toxins, and we reiterate that suggestion here. Full details of this classification code can be found in the relevant reference (Kusumi et al., 2011), but we provide a brief summary of some of the most relevant points below:

- Gene symbols should be written completely in lower case and in italics, for example "*gene2"*.
- Punctuation (dashes, periods, slashes) should not be used unless they are part of a human or mouse gene symbol.

- Whenever orthology can be assigned, this should be present in the gene name, for example if the human gene symbol is "*GENE2"* then the reptile gene symbol would be "*gene2"*.
- If an orthologous gene cannot be identified in any currently sequenced genome, a novel name may be selected by the investigators.
- Gene symbols should not start with letters to indicate genus/species.
- Gene duplicates should be assigned the suffix "a" or "b" to indicate them as being paralogs, e.g. *gene2a* and *gene2b*.

In addition, we would suggest that where toxins are known only from peptide or protein sequences, without accompanying characterized gene sequences, they should be named on the basis of similarity to existing toxins using the suffix "–like" or be acknowledged as uncharacterized toxins. We should not shy away from acknowledging uncertainty, nor from presenting challenges to future researchers. That being said, it seems likely that the availability of genomic and transcriptomic data from an ever-growing range of species (see next section) will go at least some way to facilitating the identification of toxins in proteomic studies.

With the recent publication of the whole genome sequences of two species of snake (the Burmese python *Python molurus bivittatus* and the king cobra, *Ophiophagus hannah* (Castoe et al., 2013;Vonk et al., 2013)), ongoing projects for several more (Castoe et al., 2011) and increasing amounts of transcriptomic data becoming available for a wide variety of species, it is now more important than ever that these data are made as easily accessible, understandable and useable as possible. Whilst we appreciate the historical nature of the names of many snake venom toxins and do not argue that their sometimes rich history should be neglected, we must also consider the needs of the present and the future. The field of snake venom research should not distance itself from the rest of biology via the continued use of non-standardized nomenclature. If we want to facilitate collaboration with those from other research fields it is

important that we all speak the same language – only then can we work together to better understand the origins and evolution of snake venom; its composition and function and its possible utility in the development of novel therapeutics.

We note that Toxicon, unlike many other journals, does not currently suggest or enforce a standard nomenclature of genes reported in its papers. We therefore respectfully suggest that the Editors of this journal are well-placed to lead the way in the acquisition and development of a standardized nomenclatural system for snake venom toxins (and, indeed, for the toxins of other venomous animals). After all, what is the point of being an Editor if you can't play God once in a while?

Adam D Hargreaves and John F Mulley
School of Biological Sciences, Bangor University, Brambell Building, Deiniol Road, Bangor, Gwynedd, LL57 2UW, United Kingdom

References

Card, D.C., Schield, D.R., Reyes-Velasco, J., Adams, R.H., Mackessy, S.P., Castoe, T.A. The genome of the Prairie Rattlesnake (*Crotalus viridis viridis*). Conference abstract from Biology of the pitvipers symposium 2, 4th-7th June 2014. Tulsa, OK, USA.

Aird, S.D., 1990. Call for an IST nomenclatural committee. Toxicon 28, 136-137.

Castoe, T.A., Bronikowski, A.M., Brodie, E.D.,3rd, Edwards, S.V., Pfrender, M.E., Shapiro, M.D., Pollock, D.D., Warren, W.C., 2011. A proposal to sequence the genome of a garter snake (*Thamnophis sirtalis*). Stand. Genomic Sci. 4, 257-270.

126 Castoe, T.A., de Koning, A.P., Hall, K.T., Card, D.C., Schield, D.R., Fujita, M.K., Ruggiero, R.P.,
127 Degner, J.F., Daza, J.M., Gu, W., Reyes-Velasco, J., Shaney, K.J., Castoe, J.M., Fox, S.E.,
128 Poole, A.W., Polanco, D., Dobry, J., Vandewege, M.W., Li, Q., Schott, R.K., Kapusta, A., Minx,
129 P., Feschotte, C., Uetz, P., Ray, D.A., Hoffmann, F.G., Bogden, R., Smith, E.N., Chang, B.S.,
130 Vonk, F.J., Casewell, N.R., Henkel, C.V., Richardson, M.K., Mackessy, S.P., Bronikowski, A.M.,
131 Yandell, M., Warren, W.C., Secor, S.M., Pollock, D.D., 2013. The Burmese python genome
132 reveals the molecular basis for extreme adaptation in snakes. Proc. Natl. Acad. Sci. U. S. A.
133 110, 20645-20650.

134 Hargreaves, A.D., Swain, M.T., Hegarty, M.J., Logan, D.W. and Mulley, J.F., 2014a Restriction
135 and recruitment – gene duplication and the origin and evolution of snake venom toxins. Genome
136 Biol. Evol. 6, 2088-2095.

137 Hargreaves, A.D., Swain, M.T., Logan, D.W., Mulley, J.F., 2014b. Testing the Toxicofera:
138 comparative reptile transcriptomics casts doubt on the single, early evolution of the reptile
139 venom system. bioRxiv doi: http://dx.doi.org/10.1101/006031

140 Kaiser, I.I., 1990. Toxin nomenclature. Toxicon 28, 137-138.

141 King, G.F., Gentz, M.C., Escoubas, P., Nicholson, G.M., 2008. A rational nomenclature for
142 naming peptide toxins from spiders and other venomous animals. Toxicon 52, 264-276.

143 Kumar, T., Radha, R., Rambhav, S., 1990. What's in a name? Toxicon 28, 135-136.

144 Kusumi, K., Kulathinal, R.J., Abzhanov, A., Boissinot, S., Crawford, N.G., Faircloth, B.C., Glenn,
145 T.C., Janes, D.E., Losos, J.B., Menke, D.B., Poe, S., Sanger, T.J., Schneider, C.J., Stapley, J.,
146 Wade, J., Wilson-Rawls, J., 2011. Developing a community-based genetic nomenclature for
147 anole lizards. BMC Genomics 12, 554-2164-12-554.

148 Oliveira, J.S., Fuentes-Silva, D., King, G.F., 2012. Development of a rational nomenclature for
149 naming peptide and protein toxins from sea anemones. Toxicon 60, 539-550.

150 Pung, Y.F., Wong, P.T., Kumar, P.P., Hodgson, W.C., Kini, R.M., 2005. Ohanin, a novel protein
151 from king cobra venom, induces hypolocomotion and hyperalgesia in mice. J. Biol. Chem. 280,
152 13137-13147.

153 Rao, V.S., Swarup, S., Kini, R.M., 2004. The catalytic subunit of pseutarin C, a group C
154 prothrombin activator from the venom of Pseudonaja textilis, is structurally similar to mammalian
155 blood coagulation factor Xa. Thromb. Haemost. 92, 509-521.

156 Reza, M., Swarup, S., Kini, R., 2007. Structure of two genes encoding parallel prothrombin
157 activators in Tropidechis carinatus snake: gene duplication and recruitment of factor X gene to
158 the venom gland. J. Thromb. Haemost. 5, 117-126.

159 Shows, T., Alper, C., Bootsma, D., Dorf, M., Douglas, T., Huisman, T., Kit, S., Klinger, H.,
160 Kozak, C., Lalley, P., 1979. International system for human gene nomenclature (1979) ISGN
161 (1979). Cytogenet. Genome Res. 25, 96-116.

162 Tytgat, J., Chandy, K.G., Garcia, M.L., Gutman, G.A., Martin-Eauclaire, M., van der Walt, Jurg
163 J, Possani, L.D., 1999. A unified nomenclature for short-chain peptides isolated from scorpion
164 venoms: α-KTx molecular subfamilies. Trends Pharmacol. Sci. 20, 444-447.

165 Undheim, E.A.B, Jones, A., Clauser, K.R., Holland, J.W., Pineda, S.S., King, G.F., Fry, B.G.
166 2014. Clawing through Evolution: Toxin Diversification and Convergence in the Ancient Lineage
167 Chilopoda (Centipedes). Mol Biol Evol 31, 2124-2148.

168 Vonk, F.J., Casewell, N.R., Henkel, C.V., Heimberg, A.M., Jansen, H.J., McCleary, R.J.,
169 Kerkkamp, H.M., Vos, R.A., Guerreiro, I., Calvete, J.J., Wuster, W., Woods, A.E., Logan, J.M.,

Harrison, R.A., Castoe, T.A., de Koning, A.P., Pollock, D.D., Yandell, M., Calderon, D., Renjifo, C., Currier, R.B., Salgado, D., Pla, D., Sanz, L., Hyder, A.S., Ribeiro, J.M., Arntzen, J.W., van den Thillart, G.E., Boetzer, M., Pirovano, W., Dirks, R.P., Spaink, H.P., Duboule, D., McGlinn, E., Kini, R.M., Richardson, M.K., 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proc. Natl. Acad. Sci. U. S. A. 110, 20651-20656.

Yamazaki, Y., Hyodo, F., Morita, T., 2003. Wide distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins. Arch. Biochem. Biophys. 412, 133-141.

Yamazaki, Y., Matsunaga, Y., Tokunaga, Y., Obayashi, S., Saito, M., Morita, T., 2009. Snake venom Vascular Endothelial Growth Factors (VEGF-Fs) exclusively vary their structures and functions among species. J. Biol. Chem. 284, 9885-9891.