Appendix 2-9

Chapter 2.

Exploiting genomic data to identify proteins involved in abalone reproduction.

<u>Omar Mendoza-Porras</u>; Natasha A. Botwright; Sean M. McWilliams; Mathew T. Cook; James O. Harris; Gene Wijffels and Michelle L. Colgrave

"Sequence alignment of human, oyster and snail SPAG6. The peptides in colours were detected in Protein pilot and mapped to SPAG6. The regions that contain the armadillo repeats in human SPAG6 [UniProtKB: O75602] are given. Additionally, in coloured boxes the location of each peptide along the paired sequences is shown". Alignment was conducted at http://www.ebi.ac.uk/Tools/msa/clustalo/.

Peptides identified in male abalone gonad:

- 1. YYSPGYSDALLER contained in Human armadillo repeat region 459-509
- 2. LANYNDDLAEAVVK partially contained in Human armadillo repeats 1 and 2 (regions: 31-70 and 73-112, respectively)
- 3. GDILPQLVYSLAEQNR (peptide with highest intensity) Contained in Human armadillo repeat 2 (region 73-112) with an amino acid substitution (G1C).
- 4. **IAASALSDICK** Contained in Human armadillo repeat 4 (region 157-196) with an amino acid substitution (C10A)
- 5. AVAVANVLPK (found in Pacific oyster *Crassostrea gigas* SPAG6 as a tryptic peptide)

sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	MSQRQVLQVFEQYQKARTQFVQMVAELATRPQNIETLQNAGVMSLLRTLLLDVVPTIQQT MSQRQVLQVFEQYQKARTGFVQTVAELASRPQNIETLQNAGVMSLLRPLLLDIVPTIQQT
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	AALALGE LANYNDDLAEAVVK CDILPQLVYSLAEQNR YKKAAAFVLRAVGKHSPQLAQA AALALGE LANYNDDLAEAVVK GDILPQLVYSLAEQNR YKKAAAFVLRAVAKHSPQLAQS
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	IVDCGALDTLVICLEDFDPGVKEAAAWALRYIARHNAELSQAVVDAGAVPLLVLCIQEPE VVDCGALDALVICLEEFDPGVKESAAWALGYIARHNAELAQAVVDAGAVPLLVLCIQEPE
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	IALKIIAASALSDIAK-SPELAQTVVDAGAVAHLAQMILNPDAKLKHQILSALSQVSKHS LSLKIIAASALSDICK-SPELAQTVVDAGAIAHLAQMILNPDAKLKRQVFSALSQISKHS
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	VDLAEMVVEAEIFPVVLTCLKDKDEYVKKNASTLIREIAKHTPELSQLVVNAGGVAAVID VDLAEMVVEAEIFPAVLTCLKDPDEYVKKNVATLIREIAKHTPELAQLIVNAGGVAAVVD
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	CIGSCKGNTRLPGIMMLGYVAAHSENLAMAVIISKGVPQLSVCLSEEPEDHIKAAAAWAL YVGESKGNIRLPGIMMLGYVAAHSENLAMAVIVSKGVVQLAIALSEEPEDHIQAAAAWAL
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	GQIGRHTPEHAIAVAVTNTLPV.LSLYMSTESSEDLQVKSKKAIKNILQKCTYLPALEPF GQIGRHTPEHAIAVAVANVLPK.LQCYLRADASEDLQTKSKKALKNILQKCVHLPALEPL
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	LYDAPPNILKHVVGQFSKVLPHDSKARRLFVTSGGLKKVQEIKAEPGSLLQEYINSINSC LHDAPPNILKHVVGQFSKVLPHDSKARRLFVTSAGLKKIQEIKAEPGSALAEYINTINNC LHDAPPNILKHVVGQFSKVLPHDSKARRLFVTSSGLKKIQEIKADPGSALAEYINTINNC *:**********************************
sp 075602 SPAG6_HUMAN gi 405950834 gb EKC18795.1 OM-I.obsoleta	YPEEIVFYYSPGYSDTLLQFVDSYQPLNN YPEEIVFYYSPGYSDTLLEFVEQYQPSK- YPEEIVFYYSPGYSDALLEFVEQFQPKA- ******