Supplemental Materials

Low mitogenomic diversity in the sperm whale (*Physeter macrocephalus*): historical demography and phylogeography reveal rapid global expansion.

Phillip A. Morin, Andrew D. Foote, C. Scott Baker, Brittany L. Hancock-Hanser, Kristen Kaschner, Bruce R. Mate, Sarah L. Mesnick, Victoria L. Pease, Hal Whitehead, Patricia E. Rosel, Alana Alexander

Contents:

Table S1: Sample information 2

Table S2. Parameters for mtDNA demographic analysis. 7

Table S3. Cetacean mitogenome accession numbers 8

Table S4. BEAST priors, cetaceans 10

Table S5. Haplotype IDs and accession numbers 11

Table S6. Haplotypic diversity 13

Table S7. Control region haplotypes 15

Table S8. AquaMaps suitable habitat 19

Table S9. PAML analysis of selection in mitogenomes 21

Table S10. Mitogenome amino acids under selection 22

Figure S1. PSMC plots and analysis methods. 25

Figure S2: Subsampling BioGeoBears clades. 30

Figure S3. AquaMaps habitat models 31

Figure S4. Sites under positive selection 35

Figure S5. BEAST Phase I phylogenetic tree. 43

Figure S6. Skyline plots for ocean basins. 44

Figure S7. Mitogenome vs. control region population structure. 45

Supplemental methods 46

Supplemental Materials References 49

# Table S1: Sample information

SWFSC LABID = internal sample identification number. No. Ns = the number of Ns (missing or ambiguous nucleotide data). Sample rows are colored by region.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SWFSC LABID | region | Mitogenome Haplotype | No. Ns | Name in Figure 2 | Other ID | Latitude | Longitude | SEX | previously published Accession No. |
| 17101 | Indian Ocean | mtGen33 | 2 | mt33\_Maldives | MALDIVES000427.01 | 4.4318 | 73.5293 | M |  |
|  | Indian Ocean | mtGen54 | 0 |  | SEY220012029-053-1 |  |  |  | Warren et al. 2017 |
|  | Indian Ocean | mtGen54 | 0 |  | SEY420021031-063-1 |  |  |  | Warren et al. 2017 |
| 78828 | Mediterranean | mtGen16 | 1 | mt16\_Med\_(3) | BP07-02 | 43.516666 | 7.283333 | U |  |
| 91151 | Mediterranean | mtGen16 | 1 | mt16\_Med\_(3) | Pm20080819-A | 42.883333 | 7.05 | U |  |
| 91154 | Mediterranean | mtGen16 | 5 | mt16\_Med\_(3) | Pm20090610A | 42.55 | 5.15 | U |  |
| 78834 | Mediterranean | mtGen75 | 20 | mt75\_Med | PMAC060321 | 43.433333 | 4.416666 | M |  |
| 17322 | N.E. Atlantic | mtGen02 | 3 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | NOR990802.2445 | 69.433333 | 15.766666 | M |  |
| 17320 | N.E. Atlantic | mtGen03 | 1 | mt03\_Watl\_NEAtl\_Npac\_(10) | NOR990731.2245 | 72.3 | 15.666666 | M |  |
| 7579 | N.E. Pacific | mtGen01 | 2 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | F |  |
| 7582 | N.E. Pacific | mtGen01 | 4 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | M |  |
| 7584 | N.E. Pacific | mtGen01 | 2 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | F |  |
| 7586 | N.E. Pacific | mtGen01 | 3 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | M |  |
| 7589 | N.E. Pacific | mtGen01 | 2 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | F |  |
| 7590 | N.E. Pacific | mtGen01 | 4 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 |  |  |
| 7596 | N.E. Pacific | mtGen01 | 2 | mt01\_NEPac\_(14) |  | 29.48333333 | -147.1833333 | F |  |
| 7606 | N.E. Pacific | mtGen01 | 16 | mt01\_NEPac\_(14) | MAC970603.01 | 26.033333 | -130.466666 | F |  |
| 11560 | N.E. Pacific | mtGen01 | 7 | mt01\_NEPac\_(14) | DSJ980809.06 | 27.166666 | -111.616666 | F |  |
| 11561 | N.E. Pacific | mtGen01 | 7 | mt01\_NEPac\_(14) | DSJ980809.07 | 27.166666 | -111.616666 | F |  |
| 13372 | N.E. Pacific | mtGen01 | 8 | mt01\_NEPac\_(14) | LG3 | -0.283333 | -90.283333 | F |  |
| 13373 | N.E. Pacific | mtGen01 | 0 | mt01\_NEPac\_(14) |  | -0.283333333 | -90.28333333 |  |  |
| 13375 | N.E. Pacific | mtGen01 | 9 | mt01\_NEPac\_(14) | LG6 | -0.266666 | -90.316666 | F |  |
| 13378 | N.E. Pacific | mtGen01 | 0 | mt01\_NEPac\_(14) |  | -0.066666667 | -90.91666667 |  |  |
| 52892 | N.E. Pacific | mtGen03 | 5 | mt03\_Watl\_NEAtl\_Npac\_(10) | AL0525-01 | 57.116666 | -136.25 | M |  |
| 7581 | N.E. Pacific | mtGen06 | 1 | mt06\_NPac\_(6) | MAC970527.04 | 29.483333 | -147.183333 | F |  |
| 7585 | N.E. Pacific | mtGen06 | 2 | mt06\_NPac\_(6) | MAC970527.08 | 29.483333 | -147.183333 | F |  |
| 7587 | N.E. Pacific | mtGen06 | 4 | mt06\_NPac\_(6) |  | 29.48333333 | -147.1833333 |  |  |
| 7588 | N.E. Pacific | mtGen06 | 1 | mt06\_NPac\_(6) |  | 29.48333333 | -147.1833333 |  |  |
| 7593 | N.E. Pacific | mtGen06 | 3 | mt06\_NPac\_(6) |  | 29.48333333 | -147.1833333 |  |  |
| 7595 | N.E. Pacific | mtGen06 | 3 | mt06\_NPac\_(6) |  | 29.48333333 | -147.1833333 |  |  |
| 4634 | N.E. Pacific | mtGen07 | 9 | mt07\_NPac\_(5) | MAC950923.02 | 27.4 | -111.616666 | F |  |
| 7644 | N.E. Pacific | mtGen07 | 1 | mt07\_NPac\_(5) | C74 | 37.116666 | -122.416666 | M |  |
| 30472 | N.E. Pacific | mtGen07 | 4 | mt07\_NPac\_(5) | DSJ021104.05 | 24.233333 | -175.45 | F |  |
| 51416 | N.E. Pacific | mtGen07 | 2 | mt07\_NPac\_(5) | MAC050710.02 | 38.816666 | -124.4 | F |  |
| 51417 | N.E. Pacific | mtGen07 | 5 | mt07\_NPac\_(5) | MAC050710.03 | 38.816666 | -124.4 | F |  |
| 30466 | N.E. Pacific | mtGen10 | 4 | mt10\_HI\_Mex\_Chile\_(4) | DSJ021030.01 | 20.6 | -160.916666 | F |  |
| 30467 | N.E. Pacific | mtGen10 | 6 | mt10\_HI\_Mex\_Chile\_(4) | DSJ021030.02 | 20.6 | -160.916666 | F |  |
| 66728 | N.E. Pacific | mtGen10 | 6 | mt10\_HI\_Mex\_Chile\_(4) | DSJ060811.01 | 26.466666 | -111.066666 | F |  |
| 11574 | N.E. Pacific | mtGen11 | 3 | mt11\_Mex\_Rus\_ChathamIs\_(3) | DSJ980811.08 | 24.533333 | -109.566666 | M |  |
| 7580 | N.E. Pacific | mtGen12 | 4 | mt12\_Npac\_(3) | MAC970527.03 | 29.48333333 | -147.1833333 | F |  |
| 7594 | N.E. Pacific | mtGen12 | 6 | mt12\_Npac\_(3) |  | 29.48333333 | -147.1833333 |  |  |
| 7597 | N.E. Pacific | mtGen12 | 5 | mt12\_Npac\_(3) |  | 29.48333333 | -147.1833333 |  |  |
| 7607 | N.E. Pacific | mtGen13 | 6 | mt13\_Npac\_HI\_(3) | MAC970603.02 | 26.033333 | -130.466666 | F |  |
| 61439 | N.E. Pacific | mtGen13 | 5 | mt13\_Npac\_HI\_(3) | RWB2006DEC06.01 | 19.633333 | -156.333333 | F |  |
| 61440 | N.E. Pacific | mtGen13 | 10 | mt13\_Npac\_HI\_(3) | RWB2006DEC06.02 | 19.616666 | -156.3 | F |  |
| 9570 | N.E. Pacific | mtGen14 | 5 | mt14\_GoC\_(3) | DSJ970810.01 | 27.116666 | -111.416666 | F |  |
| 9571 | N.E. Pacific | mtGen14 | 6 | mt14\_GoC\_(3) | DSJ970810.02 | 27.116666 | -111.416666 | F |  |
| 9572 | N.E. Pacific | mtGen14 | 9 | mt14\_GoC\_(3) | DSJ970810.03 | 27.116666 | -111.416666 | F |  |
| 11571 | N.E. Pacific | mtGen15 | 8 | mt15\_Mex\_(3) | DSJ980811.05 | 24.533333 | -109.566666 | F |  |
| 11572 | N.E. Pacific | mtGen15 | 11 | mt15\_Mex\_(3) | DSJ980811.06 | 24.533333 | -109.566666 | F |  |
| 11573 | N.E. Pacific | mtGen15 | 9 | mt15\_Mex\_(3) | DSJ980811.07 | 24.533333 | -109.566666 | F |  |
| 13376 | N.E. Pacific | mtGen23 | 2 | mt23\_Galapagos\_(2) |  | -0.183333333 | -90.33333333 |  |  |
| 13379 | N.E. Pacific | mtGen23 | 1 | mt23\_Galapagos\_(2) |  | 0 | -90 |  |  |
|  | N.E. Pacific | mtGen24 | 0 |  | SC991024-177-1 |  |  |  | Warren et al. 2017 |
| 15872 | N.E. Pacific | mtGen24 | 8 | mt24\_Mex\_(2) | DSJ990804.04 | 22.833333 | -110.466666 | F |  |
| 15877 | N.E. Pacific | mtGen24 | 1 | mt24\_Mex\_(2) | DSJ990804.09 | 22.833333 | -110.466666 | F |  |
| 15873 | N.E. Pacific | mtGen25 | 7 | mt25\_Mex\_(2) | DSJ990804.05 | 22.833333 | -110.466666 | F |  |
| 15876 | N.E. Pacific | mtGen25 | 7 | mt25\_Mex\_(2) | DSJ990804.08 | 22.833333 | -110.466666 | F |  |
| 15975 | N.E. Pacific | mtGen26 | 9 | mt26\_Mex\_(2) | DSJ990815.10 | 19.183333 | -105.533333 | F |  |
| 15977 | N.E. Pacific | mtGen26 | 11 | mt26\_Mex\_(2) | DSJ990815.12 | 19.183333 | -105.533333 | F |  |
| 25355 | N.E. Pacific | mtGen27 | 5 | mt27\_ENPac\_Peru\_(2) | DSJ010731.01 | 32.916666 | -121.116666 | F |  |
| 49068 | N.E. Pacific | mtGen28 | 1 | mt28\_Palmyra\_(2) | MAC050820.05 | 6.483333 | -162.4 | F |  |
| 49069 | N.E. Pacific | mtGen28 | 3 | mt28\_Palmyra\_(2) | MAC050820.06 | 6.483333 | -162.4 | F |  |
| 55207 | N.E. Pacific | mtGen29 | 6 | mt29\_HI\_(2) | RWB2006APR24.01 | 19.766666 | -156.166666 | F |  |
| 55208 | N.E. Pacific | mtGen29 | 7 | mt29\_HI\_(2) | RWB2006APR24.02 | 19.783333 | -156.166666 | F |  |
|  | N.E. Pacific | mtGen44 | 19 | mt44\_ENPac | PmaOR001 | 44.66 | -124.166 |  | KC312619 |
| 4633 | N.E. Pacific | mtGen46 | 7 | mt46\_GoC | MAC950923.01 | 27.4 | -111.616666 | F |  |
| 7546 | N.E. Pacific | mtGen48 | 4 | mt48\_Npac | MAC970402.02 | 20.583333 | -157.483333 | F |  |
| 7604 | N.E. Pacific | mtGen49 | 7 | mt49\_Npac | MAC970531.03 | 26.766666 | -138.583333 | F |  |
| 11535 | N.E. Pacific | mtGen55 | 8 | mt55\_Mex | DSJ980804.07 | 24.15 | -112.2 | F |  |
| 13371 | N.E. Pacific | mtGen57 | 5 | mt57\_Galapagos | LG2 | -0.25 | -90.3 | F |  |
| 13968 | N.E. Pacific | mtGen58 | 6 | mt58\_ENPac | LBL99-01 | 49 | -127 | F |  |
| 15974 | N.E. Pacific | mtGen59 | 9 | mt59\_Mex | DSJ990815.09 | 19.183333 | -105.533333 | F |  |
| 25358 | N.E. Pacific | mtGen63 | 2 | mt63\_ENPac | DSJ010801.03 | 33.35 | -123.066666 | F |  |
|  | N.E. Pacific | mtGen69 | 0 |  | SC991025-181-1 |  |  |  | Warren et al. 2017 |
| 49093 | N.E. Pacific | mtGen72 | 4 | mt72\_Palmyra | MAC050829.01 | 5.7 | -163.916666 | F |  |
| 51419 | N.E. Pacific | mtGen73 | 0 | mt73\_NEPac | MAC050710.05 | 38.816666 | -124.4 | F |  |
| 7592 | N.E. Pacific | mtGen80 | 1 | mt80\_NPac |  | 29.48333333 | -147.1833333 | M |  |
| 25357 | N.E. Pacific | mtGen81 | 6 | mt81\_ENPac | DSJ010801.02 | 33.283333 | -122.766666 | F |  |
| 70558 | N.W. Pacific | mtGen11 | 2 | mt11\_Mex\_Rus\_ChathamIs\_(3) | RU-SpWhale-20050714 | 46.4 | 150.716666 | M |  |
| 2659 | N.W. Pacific | mtGen45 | 1 | mt45\_Philippines | ZZZ0013 | 9.30895 | 123.503 | M |  |
| 4978 | N.W. Pacific | mtGen47 | 0 | mt47\_Japan | JFA-96-01-22 | 38.9 | 141.583333 | M |  |
| 35382 | N.W. Pacific | mtGen70 | 3 | mt70\_HongKong | PM03-21/07 | 22 | 114 | M |  |
| 70557 | N.W. Pacific | mtGen74 | 6 | mt74\_Russia | SpWRU-040720#1 | 48.333333 | 153.45 | M |  |
| 34661 | S.E. Pacific | mtGen10 | 3 | mt10\_HI\_Mex\_Chile\_(4) | CHILE-PM-C80/5438 | -20.47 | -71.37 | U |  |
| 16092 | S.E. Pacific | mtGen20 | 5 | mt20\_NZ\_Peru\_(2) | DSJ991106.09 | -5.466666 | -84.483333 | F |  |
| 16098 | S.E. Pacific | mtGen27 | 9 | mt27\_ENPac\_Peru\_(2) | DSJ991106.15 | -5.466666 | -84.483333 | F |  |
| 16062 | S.E. Pacific | mtGen60 | 6 | mt60\_Peru | DSJ991027.02 | -13.65 | -77.766666 | F |  |
| 16063 | S.E. Pacific | mtGen61 | 7 | mt61\_Peru | DSJ991027.03 | -13.65 | -77.766666 | F |  |
| 16091 | S.E. Pacific | mtGen62 | 7 | mt62\_Peru | DSJ991106.08 | -5.466666 | -84.483333 | F |  |
| 34636 | S.E. Pacific | mtGen64 | 10 | mt64\_Chile | CHILE-PM-C55 | -24.29 | -71.13 | U |  |
| 34657 | S.E. Pacific | mtGen65 | 15 | mt65\_Chile | CHILE-PM-C76/5429 | -21.09 | -71.37 | U |  |
| 34669 | S.E. Pacific | mtGen66 | 8 | mt66\_Chile | CHILE-PM-C88/5483 | -22.03 | -71.03 | U |  |
| 34672 | S.E. Pacific | mtGen67 | 65 | mt67\_Chile | CHILE-PM-C91/5034 | -22.37 | -70.49 | U |  |
| 34718 | S.E. Pacific | mtGen68 | 28 | mt68\_Chile | CHILE-PM-C137/5690 | -22.16 | -71.51 | U |  |
| 34783 | S.E. Pacific | mtGen69 | 27 | mt69\_Chile | CHILE-PM-C202 | -21.35 | -71.28 | U |  |
| 9663 | S.W. Pacific | mtGen04 | 3 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9664 | S.W. Pacific | mtGen04 | 1 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9665 | S.W. Pacific | mtGen04 | 1 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9666 | S.W. Pacific | mtGen04 | 5 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9667 | S.W. Pacific | mtGen04 | 1 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9668 | S.W. Pacific | mtGen04 | 2 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9669 | S.W. Pacific | mtGen04 | 5 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9670 | S.W. Pacific | mtGen04 | 1 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9671 | S.W. Pacific | mtGen04 | 4 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9672 | S.W. Pacific | mtGen04 | 3 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
| 9673 | S.W. Pacific | mtGen04 | 11 | mt04\_Tasmania\_strand\_(11) |  | -40.83333333 | 145.3 | F |  |
|  | S.W. Pacific | mtGen11 | 10 | mt11\_Mex\_Rus\_ChathamIs\_(3) | PmaNZ039 | -43.815226 | -176.619061 |  | KC312612 |
| 123943 | S.W. Pacific | mtGen18 | 0 | mt18\_NZ\_(1) | PmaNZ005 | -42.618802 | 173.469463 | M | KC312603.2 |
| 37718 | S.W. Pacific | mtGen19 | 2 | mt19\_NZ\_(2) | PM030729-23 | -19.083333 | -174.45 | U |  |
|  | S.W. Pacific | mtGen19 | 10 | mt19\_NZ\_(2) | PmaNZ015 | -35.989049 | 173.763871 |  | KC312607 |
|  | S.W. Pacific | mtGen20 | 10 | mt20\_NZ\_Peru\_(2) | PmaNZ036 | -43.815226 | -176.619061 |  | KC312611 |
| 123945 | S.W. Pacific | mtGen21 | 1 | mt21\_NZ\_(1) | PmaNZ034 | -37.651568 | 177.858678 | F | KC312610.2 |
|  | S.W. Pacific | mtGen31 | 10 | mt31\_NZ | PmaNZ032 | -36.841243 | 174.419032 |  | KC312609 |
|  | S.W. Pacific | mtGen34 | 10 | mt34\_NZ | PmaNZ008 | -39.854588 | 174.688226 |  | KC312604 |
|  | S.W. Pacific | mtGen35 | 10 | mt35\_NZ | PmaNZ010 | -39.054376 | 177.574457 |  | KC312605 |
|  | S.W. Pacific | mtGen37 | 10 | mt37\_NZ | PmaNZ016 | -36.845182 | 174.417935 |  | KC312608 |
|  | S.W. Pacific | mtGen38 | 10 | mt38\_NZ | PmaNZ045 | -34.81274781 | 173.0095131 |  | KC312613 |
|  | S.W. Pacific | mtGen39 | 10 | mt39\_NZ | PmaNZ056 | -38.050035 | 174.722093 |  | KC312614 |
|  | S.W. Pacific | mtGen40 | 10 | mt40\_NZ | PmaNZ058 | -39.062263 | 177.835932 |  | KC312615 |
|  | S.W. Pacific | mtGen41 | 19 | mt41\_NZ | PmaNZ076 | -38.573559 | 178.228339 |  | KC312616 |
|  | S.W. Pacific | mtGen42 | 19 | mt42\_NZ | PmaNZ082 | -35.506885 | 174.474826 |  | KC312617 |
|  | S.W. Pacific | mtGen43 | 10 | mt43\_NZ | PmaNZ085 | -44.00281588 | 168.508327 |  | KC312618 |
| 9599 | S.W. Pacific | mtGen50 | 2 | mt50\_Tasmania | AUS-OB-PM-001 | -42.2 | 145.216666 | F |  |
| 9603 | S.W. Pacific | mtGen51 | 5 | mt51\_Tasmania | AUS-OB-PM-005 | -42.2 | 145.216666 | F |  |
| 9687 | S.W. Pacific | mtGen52 | 12 | mt52\_Tasmania | AUS-MW-PM-014 | -40.9 | 144.683333 | M |  |
| 9693 | S.W. Pacific | mtGen53 | 1 | mt53\_Tasmania | AUS-MW-PM-020 | -40.9 | 144.683333 | M |  |
| 9708 | S.W. Pacific | mtGen54 | 7 | mt54\_Tasmania | AUS-MW-PM-035 | -40.9 | 144.683333 | M |  |
| 37723 | S.W. Pacific | mtGen71 | 5 | mt71\_NZ | PM030820-28 | -19.166666 | -174.466666 | U |  |
| 123944 | S.W. Pacific | mtGen78 | 19 | mt78\_NZ | PmaNZ013 | -41.733279 | 174.267996 | M | KC312606.2 |
| 1391 | W. Atlantic | mtGen02 | 5 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | 85-12-92-NHCW | 34.2 | -77.783333 | F |  |
| 2514 | W. Atlantic | mtGen02 | 1 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | MS-008-94 | 30.233333 | -88.733333 | F |  |
| 7434 | W. Atlantic | mtGen02 | 6 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | MSPM-9401 | 26.333333 | -80.083333 | M |  |
| 117569 | W. Atlantic | mtGen02 | 3 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pmac001 | 35.1933 | -75.73 | F |  |
| 117572 | W. Atlantic | mtGen02 | 1 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pmac007 | 33.6762 | -78.8958 | F |  |
| 117575 | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pmac010 | 41.0467 | -68.6683 | U |  |
| 117577 | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pmac056 | 40.38 | -67.65 | U |  |
| 117584 | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pmac066 | 26.3432 | -80.0699 | M |  |
| 133990 | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | 110616\_Pm2b | 24.67166667 | -77.56733333 | M |  |
|  | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pma11GMX10 | 27.97579 | -87.67041 | U |  |
|  | W. Atlantic | mtGen02 | 1 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pma11GMX11 | 27.922 | -87.6842 | U |  |
|  | W. Atlantic | mtGen02 | 0 | mt02\_NWAtl\_GoM\_NEAtl\_(13) | Pma11GMX13 | 27.9801 | -87.66265 | U |  |
|  | W. Atlantic | mtGen02 | 0 |  | GMX-SRS387925 |  |  | F | Warren et al. 2017 |
| 9832 | W. Atlantic | mtGen03 | 12 | mt03\_Watl\_NEAtl\_Npac\_(10) | PM3 | -3.716666 | -38.5 | U |  |
| 9833 | W. Atlantic | mtGen03 | 2 | mt03\_Watl\_NEAtl\_Npac\_(10) | PM4 | -4 | -38 | F |  |
| 12604 | W. Atlantic | mtGen03 | 5 | mt03\_Watl\_NEAtl\_Npac\_(10) | DEL990419.02 | 36.483333 | -74.583333 | M |  |
| 12695 | W. Atlantic | mtGen03 | 0 | mt03\_Watl\_NEAtl\_Npac\_(10) | WAM0547 | 35.4912 | -75.4776 | F |  |
| 117578 | W. Atlantic | mtGen03 | 1 | mt03\_Watl\_NEAtl\_Npac\_(10) | Pmac057 | 37.0074 | -73.7991 | U |  |
| 117580 | W. Atlantic | mtGen03 | 1 | mt03\_Watl\_NEAtl\_Npac\_(10) | Pmac061 | 35.3469 | -75.5009 | M |  |
| 117582 | W. Atlantic | mtGen03 | 1 | mt03\_Watl\_NEAtl\_Npac\_(10) | Pmac063 | 34.5873 | -76.5356 | F |  |
| 125732 | W. Atlantic | mtGen03 | 3 | mt03\_Watl\_NEAtl\_Npac\_(10) | 110506\_Pm10c | 24.65366667 | -77.502 | F |  |
| 117581 | W. Atlantic | mtGen05 | 0 | mt05\_NWatl\_GoM\_(8) | Pmac062 | 34.6227 | -76.5544 | M |  |
| 124019 | W. Atlantic | mtGen05 | 2 | mt05\_NWatl\_GoM\_(8) | 02091101PM/Pmac029 | 28.666666 | -88.983333 | F |  |
| 124020 | W. Atlantic | mtGen05 | 1 | mt05\_NWatl\_GoM\_(8) | 03071001PM/Pmac030 | 28.266666 | -89.716666 | F |  |
| 124022 | W. Atlantic | mtGen05 | 2 | mt05\_NWatl\_GoM\_(8) | 04070602PM/Pmac032 | 28.1 | -89.683333 | F |  |
|  | W. Atlantic | mtGen05 | 0 | mt05\_NWatl\_GoM\_(8) | Pma11GMX04 | 28.55545 | -88.81815 | U |  |
|  | W. Atlantic | mtGen05 | 0 | mt05\_NWatl\_GoM\_(8) | Pma11GMX05 | 28.55273 | -88.80983 | U |  |
|  | W. Atlantic | mtGen05 | 1 | mt05\_NWatl\_GoM\_(8) | Pma11GMX06 | 28.5309 | -88.79465 | U |  |
|  | W. Atlantic | mtGen05 | 1 | mt05\_NWatl\_GoM\_(8) | Pma11GMX15 | 28.52112 | -88.89601 | U |  |
| 7433 | W. Atlantic | mtGen09 | 1 | mt09\_NWAtl\_(4) | FMCC-9501 | 26.766666 | -80.033333 | F |  |
| 79844 | W. Atlantic | mtGen09 | 5 | mt09\_NWAtl\_(4) | 020118\_Pm3a | 26.15 | -77.616666 | F |  |
| 117571 | W. Atlantic | mtGen09 | 3 | mt09\_NWAtl\_(4) | Pmac003 | 40.39 | -67.2464 | M |  |
| 126049 | W. Atlantic | mtGen09 | 2 | mt09\_NWAtl\_(4) | 091030\_Pm2c | 24.69166667 | -77.48733333 | M |  |
| 117574 | W. Atlantic | mtGen17 | 0 | mt17\_NWAtl\_(3) | Pmac009 | 25.194 | -80.0851 | F |  |
| 117579 | W. Atlantic | mtGen17 | 5 | mt17\_NWAtl\_(3) | Pmac058 | 36.9267 | -74.2763 | U |  |
| 117583 | W. Atlantic | mtGen17 | 16 | mt17\_NWAtl\_(3) | Pmac064 | 32.9917 | -79.551 | F |  |
| 117570 | W. Atlantic | mtGen22 | 2 | mt22\_NWAtl\_GoM\_(2) | Pmac002 | 35.7674 | -75.5347 | F |  |
|  | W. Atlantic | mtGen22 | 0 | mt22\_NWAtl\_GoM\_(2) | Pma11GMX09 | 29.1829 | -87.1198 | U |  |
| 124021 | W. Atlantic | mtGen30 | 1 | mt30\_GoM\_(2) | 03070501PM/Pmac031 | 28.633333 | -88.75 | F |  |
| 124023 | W. Atlantic | mtGen30 | 1 | mt30\_GoM\_(2) | 05062601PM/Pmac033 | 27.3 | -93.766666 | F |  |
|  | W. Atlantic | mtGen32 | 1 | mt32\_GoM | Pma11GMX08 | 29.1714 | -87.0558 | U |  |
| 12127 | W. Atlantic | mtGen56 | 22 | mt56\_NWAtl | 97PM1A | 25.67 | -76.97 | U |  |
| 117573 | W. Atlantic | mtGen76 | 3 | mt76\_NWAtl | Pmac008 | 36.445 | -75.8178 | M |  |
| 117576 | W. Atlantic | mtGen77 | 1 | mt77\_NWAtl | Pmac055 | 40.38 | -67.65 | U |  |
| 126096 | W. Atlantic | mtGen79 | 3 | mt79\_NWAtl | 110606\_Pm1b | 26.482 | -78.12133333 | F |  |
| 116872 | W. Pacific | mtGen08 | 2 | mt08\_Marianas\_(5) | PIC130723.01B | 15.35516909 | 145.5555419 | F |  |
| 116875 | W. Pacific | mtGen08 | 2 | mt08\_Marianas\_(5) | PIC130723.04B | 15.38556078 | 145.5455064 | F |  |
| 116876 | W. Pacific | mtGen08 | 1 | mt08\_Marianas\_(5) | PIC130723.05B | 15.39212247 | 145.5474597 | F |  |
| 116877 | W. Pacific | mtGen08 | 2 | mt08\_Marianas\_(5) | PIC130723.06B | 15.38568811 | 145.5477092 | F |  |
| 116878 | W. Pacific | mtGen08 | 5 | mt08\_Marianas\_(5) | PIC130723.07B | 15.40039625 | 145.5475718 | F |  |

# Table S2. Parameters for mtDNA demographic analysis.

Parameters for the skyride analysis following Alexander et al. (2013). Categories and Parameter ID names follow BEAST v1.8 designations, with cp1 through cp3 referring to codon position 1 through 3 of the protein coding regions, respectively.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Category** | **Parameter ID (prior shape)** | **Statistic** | **cp1** | **cp2** | **cp3** |
| constantSize | initialDemo.popSize | Value | 0.005 | | |
| gmrfSkyrideLikelihood | skyride.logPopSize | Dimension | 174 | | |
|  | skyride.logPopSize | Value | -5.298317367 | | |
|  | skyride.groupSize | Dimension | 174 | | |
|  | skyride.precision (gamma) | Shape | 0.001 | | |
|  | Scale | 1000 | | |
| strictClockBranchRates | clock.rate | Mean | 0.00468936 | 0.0013 | 0.02521064 |
|  | Stdev | -- | 8.91E-05 | 0.001299688 |
| gtrModel | frequencies (uniform) | Upper | 1 | 1 | 1 |
|  | Lower | 0 | 0 | 0 |
|  | ac (normal) | Mean | 0.03215 | 0.1416 | 0.023048 |
|  | Stdev | 0.004896 | 0.030361 | 0.002396 |
|  | ag (normal) | Mean | 0.4227 | 1.1444 | 2.0215 |
|  | Stdev | 0.062066 | 0.18088 | 0.2865 |
|  | at (normal) | Mean | 0.057532 | 0.073491 | 0.032688 |
|  | Stdev | 0.007044 | 0.021644 | 0.003571 |
|  | cg (normal) | Mean | 0.009186 | 0.1487 | 0.042182 |
|  | Stdev | 0.003449 | 0.039554 | 0.012722 |
|  | gt (normal) | Value | 0.034881 | 0.025768 | 0.07722 |
|  | Lower | 0.007155 | 0.009049 | 0.023655 |
| siteModel | gammaCategories |  | 4 | 4 | 4 |
|  | alpha (normal) | Mean | 0.6801 | 0.5291 | 3.2493 |
|  | Stdev | 0.15 | 0.115 | 0.269 |
|  | pInv (normal) | Mean | 0.4892 | 0.6861 | 0.015108 |
|  | Stdev | 0.079503 | 0.058858 | 0.004146 |

# Table S3. Cetacean mitogenome accession numbers

Genbank accession numbers and references for the cetacean mitogenomes used in the selection analyses.

|  |  |  |
| --- | --- | --- |
| **Species** | **Accession #** | **Reference** |
| *Steno bredanensis* | JF339982 | Vilstrup et al. (2011) |
| *Balaena mysticetus* | NC\_005268 | Arnason et al. (2004) |
| *Balaenoptera acutorostrata* | NC\_005271 | Arnason et al. (2004) |
| *Balaenoptera bonaerensis* | NC\_006926 | Sasaki et al. (2005) |
| *Balaenoptera borealis* | NC\_006929 | Sasaki et al. (2005) |
| *Balaenoptera brydei* | NC\_006928 | Sasaki et al. (2005) |
| *Balaenoptera edeni* | NC\_007938 | Sasaki et al. (2006) |
| *Balaenoptera musculus* | NC\_001601 | Arnason and Gullberg (1993) |
| *Balaenoptera omurai* | NC\_007937 | Sasaki et al. (2006) |
| *Balaenoptera physalus* | NC\_001321 | Arnason et al. (1991) |
| *Berardius bairdii* | NC\_005274 | Arnason et al. (2004) |
| *Caperea marginata* | NC\_005269 | Arnason et al. (2004) |
| *Cephalorhynchus commersonii* | KC312623 | Alexander et al. (2013) |
| *Cephalorhynchus eutropia* | KC312624 | Alexander et al. (2013) |
| *Cephalorhynchus heavisidii* | NC\_020696 | Hassanin et al. (2012) |
| *Cephalorhynchus hectori hectori* | KC312626 | Alexander et al. (2013) |
| *Cephalorhynchus hectori maui* | KC312627 | Alexander et al. (2013) |
| *Delphinus capensis* | NC\_012061 | Xiong et al. (2009) |
| *Delphinus delphis* | KC312628 | Alexander et al. (2013) |
| *Eschrichtius robustus* | NC\_005270 | Arnason et al. (2004) |
| *Eubalaena australis* | NC\_006930 | Sasaki et al. (2005) |
| *Eubalaena japonica* | NC\_006931 | Sasaki et al. (2005) |
| *Feresa attenuata* | NC\_019588 | Vilstrup et al. (2011) |
| *Globicephala melas* | NC\_019578 | Morin et al. (2010) |
| *Globicephala melas* | NC\_019441 | Vilstrup et al. (2011) |
| *Grampus griseus* | NC\_012062 | Xiong et al. (2009) |
| *Hyperoodon ampullatus* | NC\_005273 | Arnason et al. (2004) |
| *Inia geoffrensis* | NC\_005276 | Arnason et al. (2004) |
| *Kogia breviceps* | NC\_005272 | Arnason et al. (2004) |
| *Lagenorhynchus albirostris* | NC\_005278 | Arnason et al. (2004) |
| *Lagenorhynchus australis* | KC312630 | Alexander et al. (2013) |
| *Lagenorhynchus cruciger* | KC312620 | Alexander et al. (2013) |
| *Lagenorhynchus obscurus* | KC312621 | Alexander et al. (2013) |
| *Lipotes vexillifer* | NC\_007629 | Yan et al. (2005) |
| *Megaptera novaeangliae* | NC\_006927 | Sasaki et al. (2005) |
| *Mesoplodon densirostris* | NC\_021974 | Morin et al. (2012) |
| *Mesoplodon europaeus* | NC\_021434 | Morin et al. (2012) |
| *Mesoplodon ginkgodens* | NC\_027593 | Yao et al. (2016) |
| *Mesoplodon grahi* | NC\_023830 | Thompson et al. (2016) |
| *Monodon monoceros* | NC\_005279 | Arnason et al. (2004) |
| *Neophocaena asiaeorientalis* | NC\_026456 | Liu et al. (2016) |
| *Neophocaena phocaenoides* | NC\_021461 | Xu et al. (2013) |
| *Orcaella brevirostris* | NC\_019590 | Vilstrup et al. (2011) |
| *Orcaella heinsohni* | NC\_019591 | Vilstrup et al. (2011) |
| *Orcinus orca* | NC\_014682 | Foote et al. (2011) |
| *Peponocephala electra* | NC\_019589 | Vilstrup et al. (2011) |
| *Phocoena phocoena* | NC\_005280 | Arnason et al. (2004) |
| *Physeter macrocephalus* (MTGEN01) | KU891329 | This study |
| *Platanista minor* | NC\_005275 | Arnason et al. (2004) |
| *Pontoporia blainvillei* | NC\_005277 | Arnason et al. (2004) |
| *Pseudorca crassidens* | NC\_019577 | Morin et al. (2010) |
| *Sousa chinensis* | NC\_012057 | Xiong et al. (2009) |
| *Stenella attenuata* | NC\_012051 | Xiong et al. (2009) |
| *Stenella coeruleoalba* | NC\_012053 | Xiong et al. (2009) |
| *Tursiops aduncus* | NC\_012058 | Xiong et al. (2009) |
| *Tursiops australis* | NC\_022805 | Moura et al. (2013) |
| *Tursiops truncatus* | NC\_012059 | Xiong et al. (2009) |
| *Ziphius cavirostris* | NC\_021435 | Morin et al. (2012) |

# Table S4. BEAST priors, cetaceans

Prior values for BEAST phylogenetic analysis of cetacean mitogenomes (see Table S3 for sequences used in analysis).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Category** | **Parameter ID (prior shape)** | **Statistic** | **cp1** | **cp2** | **cp3** | **Notes** |
| Yule Model | birthRate (uniform) | Upper | 0 | | |  |
|  |  | Lower | ∞ | | |  |
| RelaxedClock | LogNormalDistributionModel | Upper | 1 | | |  |
|  |  | Lower | 0 | | |  |
|  | ucldMean (uniform) | Upper | 0.486972\* | 0.1512 | 2.5812 | \* Point value for cp1 as other partitions estimated relative to this partition |
|  |  | Lower | -- | 0.1188 | 2.6568 |  |
|  | ucldStdev (gamma) | Alpha | 0.5396 | | |  |
|  |  | Beta | 0.3819 | | |  |
| gtrModel | frequencies (uniform) | Upper | 1 | 1 | 1 |  |
|  |  | Lower | 0 | 0 | 0 |  |
|  | ac | Value | 0.03215 | 0.1416 | 0.023048 |  |
|  | ag | Value | 0.4227 | 1.1444 | 2.0215 |  |
|  | at | Value | 0.057532 | 0.073491 | 0.032688 |  |
|  | cg | Value | 0.009186 | 0.1487 | 0.042182 |  |
|  | gt | Value | 0.034881 | 0.025768 | 0.07722 |  |
|  | gammaCategories | Value | 4 | 4 | 4 |  |
|  | Shape | Value | 0.6801 | 0.5291 | 3.2493 |  |
|  | Proportion Inavariant | Value | 0.4892 | 0.6861 | 0.015108 |  |

# Table S5. Haplotype IDs and accession numbers

Mitogenome haplotype IDs, number of samples per haplotype, and accession numbers. Three of the unique haplotypes could not be unambiguously assigned to any other existing haplotype, due to Ns at sites that were variable across the aligned mitogenome haplotypes. In all cases (haplotypes mt56, mt69, mt75) the haplotype was found in only one individual and was identical at all called nucleotides to another haplotype, but we considered them to be different haplotypes to avoid potentially conflating unique haplotypes. There is no haplotype mtGen36.

|  |  |  |  |
| --- | --- | --- | --- |
| Haplotype ID | No. Samples | Accession No. | Source |
| mtGen01 | 14 | KU891329 | This study |
| mtGen02 | 13 | KU891330 | This study |
| mtGen03 | 10 | KU891331 | This study |
| mtGen04 | 11 | KU891332 | This study |
| mtGen05 | 8 | KU891333 | This study |
| mtGen06 | 6 | KU891334 | This study |
| mtGen07 | 5 | KU891335 | This study |
| mtGen08 | 5 | KU891336 | This study |
| mtGen09 | 4 | KU891337 | This study |
| mtGen10 | 4 | KU891338 | This study |
| mtGen11 | 3 | KC312612, KU891339 | Alexander et al. 2013, This study |
| mtGen12 | 3 | KU891340 | This study |
| mtGen13 | 3 | KU891341 | This study |
| mtGen14 | 3 | KU891342 | This study |
| mtGen15 | 3 | KU891343 | This study |
| mtGen16 | 3 | KU891344 | This study |
| mtGen17 | 3 | KU891345 | This study |
| mtGen18 | 1 | KC312603.2 | Alexander et al. 2013 |
| mtGen19 | 2 | KC312607, KU891346 | Alexander et al. 2013, This study |
| mtGen20 | 2 | KC312611, KU891347 | Alexander et al. 2013, This study |
| mtGen21 | 1 | KC312610.2 | Alexander et al. 2013 |
| mtGen22 | 2 | KU891348 | This study |
| mtGen23 | 2 | KU891349 | This study |
| mtGen24 | 2 | KU891350 | This study |
| mtGen25 | 2 | KU891351 | This study |
| mtGen26 | 2 | KU891352 | This study |
| mtGen27 | 2 | KU891353 | This study |
| mtGen28 | 2 | KU891354 | This study |
| mtGen29 | 2 | KU891355 | This study |
| mtGen30 | 2 | KU891356 | This study |
| mtGen31 | 1 | KC312609 | Alexander et al. 2013 |
| mtGen32 | 1 | KU891357 | This study |
| mtGen33 | 1 | KU891358 | This study |
| mtGen34 | 1 | KC312604 | Alexander et al. 2013 |
| mtGen35 | 1 | KC312605 | Alexander et al. 2013 |
| mtGen37 | 1 | KC312608 | Alexander et al. 2013 |
| mtGen38 | 1 | KC312613 | Alexander et al. 2013 |
| mtGen39 | 1 | KC312614 | Alexander et al. 2013 |
| mtGen40 | 1 | KC312615 | Alexander et al. 2013 |
| mtGen41 | 1 | KC312616 | Alexander et al. 2013 |
| mtGen42 | 1 | KC312617 | Alexander et al. 2013 |
| mtGen43 | 1 | KC312618 | Alexander et al. 2013 |
| mtGen44 | 1 | KC312619 | Alexander et al. 2013 |
| mtGen45 | 1 | KU891359 | This study |
| mtGen46 | 1 | KU891360 | This study |
| mtGen47 | 1 | KU891361 | This study |
| mtGen48 | 1 | KU891362 | This study |
| mtGen49 | 1 | KU891363 | This study |
| mtGen50 | 1 | KU891364 | This study |
| mtGen51 | 1 | KU891365 | This study |
| mtGen52 | 1 | KU891366 | This study |
| mtGen53 | 1 | KU891367 | This study |
| mtGen54 | 1 | KU891368 | This study |
| mtGen55 | 1 | KU891369 | This study |
| mtGen56 | 1 | KU891370 | This study |
| mtGen57 | 1 | KU891371 | This study |
| mtGen58 | 1 | KU891372 | This study |
| mtGen59 | 1 | KU891373 | This study |
| mtGen60 | 1 | KU891374 | This study |
| mtGen61 | 1 | KU891375 | This study |
| mtGen62 | 1 | KU891376 | This study |
| mtGen63 | 1 | KU891377 | This study |
| mtGen64 | 1 | KU891378 | This study |
| mtGen65 | 1 | KU891379 | This study |
| mtGen66 | 1 | KU891380 | This study |
| mtGen67 | 1 | KU891381 | This study |
| mtGen68 | 1 | KU891382 | This study |
| mtGen69 | 1 | KU891383 | This study |
| mtGen70 | 1 | KU891384 | This study |
| mtGen71 | 1 | KU891385 | This study |
| mtGen72 | 1 | KU891386 | This study |
| mtGen73 | 1 | KU891387 | This study |
| mtGen74 | 1 | KU891388 | This study |
| mtGen75 | 1 | KU891389 | This study |
| mtGen76 | 1 | KU891390 | This study |
| mtGen77 | 1 | KU891391 | This study |
| mtGen78 | 1 | KC312606.2 | Alexander et al. 2013 |
| mtGen79 | 1 | KU891392 | This study |
| mtGen80 | 1 | KU891393 | This study |
| mtGen81 | 1 | KU891394 | This study |

# Table S6. Haplotypic diversity

Number of haplotypes, haplotypic diversity and nucleotide diversity by ocean basin. A) based on all samples and B) after removal of replicate haplotypes collected from the same social group to control for nonrandom sampling of close relatives. pct.gte.0 is the proportion of bootstrapped values (out of 1000) that were different from the observed value, and considered significant if it was <0.05 or >0.95.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A) |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | No. Samples | No. Haplotypes | Haplotype Diversity | Nucleotide Diversity () |  |  |  |  |
| Atlantic | 35 | 10 | 0.839 | 0.0010 |  |  |  |  |
| GoMx | 11 | 5 | 0.818 | 0.0010 |  |  |  |  |
| Mediterranean | 4 | 2 | 0.500 | 0.0000 |  |  |  |  |
| Pacific | 124 | 66 | 0.972 | 0.0008 |  |  |  |  |
| Global | 175 | 80 | 0.975 | 0.0009 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| hap diff bootstrap | |  |  |  |  |  |  |  |
| strata.1 | strata.2 | n.1 | n.2 | het.1 | het.2 | obs het diff | pct.gte.0 | significant |
| Atlantic | GoMx | 35 | 11 | 0.8387 | 0.8182 | 0.0205 | 0.7802 | no |
| Atlantic | Mediterranean | 35 | 4 | 0.8387 | 0.5000 | 0.3387 | 1.0000 | yes |
| Atlantic | Pacific | 35 | 124 | 0.8387 | 0.9723 | -0.1337 | 0.0010 | yes |
| GoMx | Mediterranean | 11 | 4 | 0.8182 | 0.5000 | 0.3182 | 0.9590 | yes |
| GoMx | Pacific | 11 | 124 | 0.8182 | 0.9723 | -0.1541 | 0.0010 | yes |
| Mediterranean | Pacific | 4 | 124 | 0.5000 | 0.9723 | -0.4723 | 0.0010 | yes |
|  |  |  |  |  |  |  |  |  |
| nuc. Div. bootstrap | |  |  |  |  |  |  |  |
| strata.1 | strata.2 | n.1 | n.2 | nuc.div.1 | nuc.div.2 | obs nuc div diff | pct.gte.0 | significant |
| Atlantic | GoMx | 35 | 11 | 0.0010 | 0.0010 | 0.0000 | 0.5485 | no |
| Atlantic | Mediterranean | 35 | 4 | 0.0010 | 0.0000 | 0.0010 | 1.0000 | yes |
| Atlantic | Pacific | 35 | 124 | 0.0010 | 0.0008 | 0.0002 | 0.9960 | yes |
| GoMx | Mediterranean | 11 | 4 | 0.0010 | 0.0000 | 0.0010 | 1.0000 | yes |
| GoM | Pacific | 11 | 124 | 0.0010 | 0.0008 | 0.0002 | 0.8511 | no |
| Mediterranean | Pacific | 4 | 124 | 0.0000 | 0.0008 | -0.0008 | 0.0010 | yes |
|  |  |  |  |  |  |  |  |  |
| B) |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | No. Samples | No. Haplotypes | Haplotype Diversity | Nucleotide Diversity () |  |  |  |  |
| Atlantic | 35 | 10 | 0.839 | 0.0010 |  |  |  |  |
| GoM | 8 | 2 | 0.786 | 0.0010 |  |  |  |  |
| Mediterranean | 4 | 2 | 0.500 | 0.0000 |  |  |  |  |
| Pacific | 87 | 66 | 0.989 | 0.0002 |  |  |  |  |
| Global | 134 | 80 | 0.978 | 0.0010 |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| hap diff bootstrap | |  |  |  |  |  |  |  |
| strata.1 | strata.2 | n.1 | n.2 | het.1 | het.2 | obs hap diff | pct.gte.0 | significant |
| Atlantic | GoMx | 35 | 8 | 0.8387 | 0.7857 | 0.0529 | 0.8102 | no |
| Atlantic | Mediterranean | 35 | 4 | 0.8387 | 0.5000 | 0.3387 | 1.0000 | yes |
| Atlantic | Pacific | 35 | 87 | 0.8387 | 0.9888 | -0.1501 | 0.0010 | yes |
| GoMx | Mediterranean | 8 | 4 | 0.7857 | 0.5000 | 0.2857 | 0.8601 | no |
| GoMx | Pacific | 8 | 87 | 0.7857 | 0.9888 | -0.2031 | 0.0010 | yes |
| Mediterranean | Pacific | 4 | 87 | 0.5000 | 0.9888 | -0.4888 | 0.0010 | yes |
|  |  |  |  |  |  |  |  |  |
| Nuc div. bootstrap | |  |  |  |  |  |  |  |
| strata.1 | strata.2 | n.1 | n.2 | nuc.div.1 | nuc.div.2 | obs nuc div diff | pct.gte.0 | significant |
| Atlantic | GoMx | 35 | 8 | 0.0010 | 0.0010 | 0.0000 | 0.5315 | no |
| Atlantic | Mediterranean | 35 | 4 | 0.0010 | 0.0000 | 0.0010 | 1.0000 | yes |
| Atlantic | Pacific | 35 | 87 | 0.0010 | 0.0008 | 0.0002 | 0.9860 | yes |
| GoMx | Mediterranean | 8 | 4 | 0.0010 | 0.0000 | 0.0010 | 1.0000 | yes |
| GoMx | Pacific | 8 | 87 | 0.0010 | 0.0008 | 0.0002 | 0.6803 | no |
| Mediterranean | Pacific | 4 | 87 | 0.0000 | 0.0008 | -0.0008 | 0.0010 | yes |

# Table S7. Control region haplotypes

1. An update of Supplementary Material 1 from Alexander et al. (2016), summarizing the variable sites that define sperm whale mtDNA CR haplotypes. Sequences from Alexander et al. (2016) start 5 bp later than those from Mesnick et al. (2011), so both numbering schemes are provided. Dashes indicate no information for a sequence at that site (no indels have been observed over the currently available CR haplotypes). Haplotypes named with reference to identity over 394, 569 and 619 bp consensus lengths. New haplotypes and/or haplotypes with additional resolution due to samples sequenced in this study are highlighted in yellow (for the given consensus length where the sequenced samples provided new information/defined new haplotypes). Mitogenome haplotype associated with CR haplotype definition given in column ‘Mitogenome Haplotype’. No new haplotypes were defined over the 394 bp consensus length.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| CR haplotypes (bp) | | | Mitogenome haplotype | Mes  Ref | 43 | 58 | 62 | 105 | 107 | 109 | 121 | 150 | 184 | 200 | 207 | 208 | 211 | 235 | 238 | 243 | 260 | 272 | 273 | 283 | 286 | 287 | 288 | 289 | 291 | 295 | 305 | 308 | 319 | 324 | 350 | 574 | 608 | 624 |
| 394 | 569 | 619 | Alx  Ref | 38 | 53 | 57 | 100 | 102 | 104 | 116 | 145 | 179 | 195 | 202 | 203 | 206 | 230 | 233 | 238 | 255 | 267 | 268 | 278 | 281 | 282 | 283 | 284 | 286 | 290 | 300 | 303 | 314 | 319 | 345 | 569 | 603 | 619 |
| A | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| A | .001 | .002 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | A |
| A | .001 | .003 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | G | G |
| A | .002 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | T | A | G |
| B | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| B | .001 | .002 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | A |
| B | .002 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | T | A | G |
| C | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | A |
| C | .001 | .002 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| C | .001 | .003 | mtGen77 |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | G | G |
| C | .002 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | T | A | G |
| D | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | C | A | A | T | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| E | .001 | .001 |  |  | T | C | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| F | .001 | .001 | mtGen26 |  | T | C | C | C | A | G | C | T | T | T | A | A | C | A | T | G | A | A | T | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| G | .NA | .NA |  |  | T | C | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | T | - | - | - |
| H | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | T | C | A | G | A | T | A | G | C | A | G | C | C | G | A | G |
| I | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | T | C | G | A | G |
| J | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| J | .002 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | T | A | G |
| K | .001 | .001 | mtGen49 |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | T | C | A | A | G | T | A | G | C | A | G | T | C | G | A | G |
| L | .001 | .001 | mtGen46 |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | A | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| M | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | G | A | A | T | A | G | C | A | G | C | C | G | A | G |
| N | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | A | C | C | G | A | G |
| N | .001 | .002 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | A | C | C | G | A | A |
| N | .002 | .NA |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | A | C | C | T | - | - |
| O | .001 | .001 |  |  | T | T | T | C | A | G | C | C | C | T | A | A | C | A | T | G | G | A | C | C | A | A | A | T | A | G | C | A | A | T | C | G | A | G |
| P | .001 | .001 | mtGen27 |  | T | T | T | C | A | G | C | C | T | T | A | A | T | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| Q | .001 | .001 |  |  | T | T | T | C | A | G | C | C | C | T | A | A | C | A | T | G | G | G | C | C | A | A | A | T | A | G | C | A | A | T | C | G | A | G |
| R | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | G | A | C | C | A | A | A | T | A | G | C | A | A | T | C | G | A | G |
| S | .001 | .001 |  |  | T | T | T | T | A | G | C | C | C | T | A | A | C | A | T | G | G | G | C | C | A | A | A | T | A | G | C | A | A | T | C | G | A | G |
| T | .NA | .NA |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | G | C | C | A | A | A | T | A | G | C | A | G | C | C | - | - | - |
| U | .NA | .NA |  |  | T | T | C | C | A | G | C | C | T | T | A | G | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | - | - | - |
| V | .NA | .NA |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | C | A | G | C | A | G | C | C | - | - | - |
| W | .001 | .001 | mtGen53 |  | T | T | C | C | A | G | C | C | T | T | A | A | T | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| X | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | G | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| Y | .001 | .001 | mtGen30 |  | T | T | T | C | A | G | C | C | T | T | G | A | C | A | T | G | A | G | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| Z | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | G | G | C | C | G | A | G |
| AA | .NA | .NA |  |  | T | C | C | C | A | G | C | T | T | T | A | A | C | G | T | G | A | A | C | C | A | A | A | T | A | G | T | A | G | C | C | - | - | - |
| BB | .001 | .001 |  |  | T | T | C | C | A | G | T | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | T | A | G |
| CC | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | T | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| DD | .001 | .001 |  |  | T | T | T | C | A | G | C | C | C | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| EE | .001 | .001 |  |  | T | T | C | C | G | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| FF | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | A | C | A | G | C | C | T | A | G |
| GG | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | C | G | A | A | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| HH | .001 | .001 |  |  | C | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | T | C | G | A | G |
| II | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | T | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| JJ | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | T | C | T | A | G |
| JJ | .002 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | A | G | C | A | G | T | C | G | A | G |
| KK | .001 | .001 |  |  | T | T | C | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | G | T | G | G | C | A | G | C | C | G | A | G |
| LL | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | A | C | T | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| MM | .001 | .001 |  |  | T | T | T | C | A | G | C | C | T | T | A | A | C | A | T | G | A | G | C | C | A | A | G | T | A | G | C | A | G | C | C | G | A | G |
| NN | .001 | .001 |  |  | T | C | C | C | A | G | C | T | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | G | C | C | G | A | G |
| OO | .001 | .001 |  |  | T | T | T | C | A | A | C | C | T | T | A | A | C | A | T | G | A | A | C | C | A | A | A | T | A | G | C | A | A | C | C | G | - | - |

1. Mitochondrial CR haplotypes associated with the unique mitogenome haplotypes found in this study. Mitogenomes highlighted in yellow match those referenced in Supplementary Material Whatever. Three mitogenome haplotypes are not depicted in the table as they had ambiguities on variable sites within the control region that defined previously described CR haplotypes, and therefore are a match to multiple CR haplotypes. These are mtGen52 (A.001.001/V.NA.NA, due to N at position 284), mtGen67 (C.001.002/C.002.001/Z.001.001, due to Ns at positions 303 and 569), and mtGen68 (A.001.001/B.001.001/I.001.001/N.001.001/O.001.001/R.001.001/ DD.001.001, due to Ns at positions 57, 179, 255, 314, 319).

|  |  |  |  |
| --- | --- | --- | --- |
| **Mitogenome haplotype** | **Control region haplotype** | | |
| **394** | **569** | **619** |
| mtGen01 | A | .001 | .001 |
| mtGen02 | B | .001 | .001 |
| mtGen03 | A | .001 | .001 |
| mtGen04 | A | .001 | .001 |
| mtGen05 | X | .001 | .001 |
| mtGen06 | D | .001 | .001 |
| mtGen07 | E | .001 | .001 |
| mtGen08 | B | .001 | .001 |
| mtGen09 | C | .001 | .002 |
| mtGen10 | J | .001 | .001 |
| mtGen11 | O | .001 | .001 |
| mtGen12 | C | .001 | .002 |
| mtGen13 | A | .001 | .001 |
| mtGen14 | C | .001 | .002 |
| mtGen15 | B | .001 | .001 |
| mtGen16 | C | .001 | .002 |
| mtGen17 | C | .001 | .002 |
| mtGen18 | B | .001 | .001 |
| mtGen19 | B | .001 | .001 |
| mtGen20 | B | .001 | .001 |
| mtGen21 | A | .001 | .003 |
| mtGen22 | B | .001 | .001 |
| mtGen23 | A | .001 | .001 |
| mtGen24 | H | .001 | .001 |
| mtGen25 | C | .001 | .002 |
| mtGen26 | F | .001 | .001 |
| mtGen27 | P | .001 | .001 |
| mtGen28 | C | .001 | .002 |
| mtGen29 | B | .001 | .001 |
| mtGen30 | Y | .001 | .001 |
| mtGen31 | B | .002 | .001 |
| mtGen32 | A | .001 | .001 |
| mtGen33 | I | .001 | .001 |
| mtGen34 | EE | .001 | .001 |
| mtGen35 | C | .001 | .002 |
| **Mitogenome haplotype** | **Control region haplotype** | | |
| **394** | **569** | **619** |
| mtGen37 | J | .001 | .001 |
| mtGen38 | B | .001 | .002 |
| mtGen39 | S | .001 | .001 |
| mtGen40 | C | .001 | .002 |
| mtGen41 | Q | .001 | .001 |
| mtGen42 | II | .001 | .001 |
| mtGen43 | N | .001 | .001 |
| mtGen44 | B | .001 | .001 |
| mtGen45 | J | .001 | .001 |
| mtGen46 | L | .001 | .001 |
| mtGen47 | A | .002 | .001 |
| mtGen48 | B | .001 | .001 |
| mtGen49 | K | .001 | .001 |
| mtGen50 | A | .001 | .001 |
| mtGen51 | C | .001 | .002 |
| mtGen53 | W | .001 | .001 |
| mtGen54 | N | .001 | .001 |
| mtGen55 | I | .001 | .001 |
| mtGen56 | C | .001 | .001 |
| mtGen57 | D | .001 | .001 |
| mtGen58 | M | .001 | .001 |
| mtGen59 | N | .001 | .001 |
| mtGen60 | J | .001 | .001 |
| mtGen61 | A | .001 | .001 |
| mtGen62 | B | .001 | .001 |
| mtGen63 | A | .001 | .001 |
| mtGen64 | I | .001 | .001 |
| mtGen65 | C | .001 | .002 |
| mtGen66 | A | .001 | .001 |
| mtGen69 | B | .001 | .001 |
| mtGen70 | B | .001 | .001 |
| mtGen71 | A | .001 | .001 |
| mtGen72 | H | .001 | .001 |
| mtGen73 | B | .001 | .001 |
| mtGen74 | A | .001 | .001 |
| mtGen75 | C | .001 | .002 |
| mtGen76 | A | .001 | .001 |
| mtGen77 | C | .001 | .003 |
| mtGen78 | J | .001 | .001 |
| mtGen79 | C | .001 | .002 |
| mtGen80 | A | .001 | .001 |
| mtGen81 | A | .002 | .001 |

# Table S8. AquaMaps suitable habitat

Area of habitat inferred from each habitat model, with proportional area of habitat by ocean relative to the present day range. Total habitat is indicated by probability threshold >0.0. Core suitable habitat is indicated by probability threshold ≥0.6 (Kaschner *et al.* 2012).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Probability** | | **Size of distribution area (km2)** | | | **Proportional size relative to current range** | | |
|  | **threshold** |  | **Pleistocene** | **Current** | **2,100** | **Pleistocene** | **Current** | **2,100** |
| **Females** | **> 0.0** |  | 267,189,230 | 301,354,579 | 304,256,196 | 0.89 | 1.00 | 1.01 |
|  | **≥ 0.6** | **(core suitable habitat)** | 168,460,794 | 188,976,254 | 229,905,931 | 0.89 | 1.00 | 1.22 |
| **Males** | **> 0.0** |  | 297,985,873 | 343,098,426 | 342,736,944 | 0.87 | 1.00 | 1.00 |
|  | **≥ 0.6** | **(core suitable habitat)** | 215,229,383 | 244,251,690 | 279,007,102 | 0.88 | 1.00 | 1.14 |
|  |  |  |  |  |  |  |  |  |
| **Females** | **> 0.0** | **Atlantic** | 62,981,425 | 76,451,384 | 77,981,675 | 0.82 | 1.00 | 1.02 |
|  |  | **IndianOcean** | 54,196,521 | 59,478,849 | 60,550,665 | 0.91 | 1.00 | 1.02 |
|  |  | **Pacific** | 148,261,934 | 162,886,690 | 165,723,857 | 0.91 | 1.00 | 1.02 |
|  |  | **Total** | **265,439,879** | **298,816,923** | **304,256,196** | 0.89 | 1.00 | 1.02 |
|  |  |  |  |  |  |  |  |  |
| **Females** | **≥ 0.6** | **Atlantic** | 26,211,502 | 51,045,127 | 48,697,030 | 0.51 | 1.00 | 0.95 |
|  |  | **IndianOcean** | 39,891,495 | 37,619,579 | 51,277,374 | 1.06 | 1.00 | 1.36 |
|  |  | **Pacific** | 102,355,279 | 100,082,878 | 129,931,527 | 1.02 | 1.00 | 1.30 |
|  |  | **Total** | **168,458,277** | **188,747,584** | **229,905,931** | 0.89 | 1.00 | 1.22 |
|  |  |  |  |  |  |  |  |  |
| **Males** | **> 0.0** | **Atlantic** | 76,847,074 | 92,195,269 | 93,326,400 | 0.83 | 1.00 | 1.01 |
|  |  | **IndianOcean** | 58,858,316 | 72,119,491 | 72,561,221 | 0.82 | 1.00 | 1.01 |
|  |  | **Pacific** | 160,528,862 | 175,706,725 | 176,849,323 | 0.91 | 1.00 | 1.01 |
|  |  | **Total** | **296,234,253** | **340,021,485** | **342,736,944** | 0.87 | 1.00 | 1.01 |
|  |  |  |  |  |  |  |  |  |
| **Males** | **≥ 0.6** | **Atlantic** | 41,141,937 | 66,756,146 | 63,473,526 | 0.62 | 1.00 | 0.95 |
|  |  | **IndianOcean** | 50,947,802 | 52,769,566 | 65,514,649 | 0.97 | 1.00 | 1.24 |
|  |  | **Pacific** | 123,137,128 | 124,497,308 | 150,018,927 | 0.99 | 1.00 | 1.20 |
|  |  | **Total** | **215,226,866** | **244,023,020** | **279,007,102** | 0.88 | 1.00 | 1.14 |

# Table S9. PAML analysis of selection in mitogenomes

Results of the PAML analysis testing for pervasive positive selection across the cetacean mitogenomes. Subscripted p values for parameter values give proportion of codons that belong to the specified omega class

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **No. of parameters** | **Log likelihood** | **Parameter values** | **LRT** | **Interpretation** |
| M0 (one ratio) | 1 | -131185.16 | ω = 0.09310 | -- | -- |
| M1a (nearly neutral) | 2 | -128397.03 | p0 = 0.89158, p1 = 0.10842 ω0 = 0.03924, ω1 = 1.00000 | D = 5576, p < 0.0001 | M1a > M0: Non-neutrality |
| M2a (positive selection) | 4 | -128397.03 | p0 = 0.89159, p1 = 0.06166, p2 = 0.04675 ω0 = 0.03924, ω1 = 1.00000, ω2 = 1.00000 | D < 0.001, p = 1.00 | M1a ≈ M2a: No evidence for positive selection |
| M7 (beta) | 2 | -127561.16 | p = 0.16094, q = 0.98197 | -- | -- |
| M8 (beta & ω) | 4 | -127457.78 | p = 0.18419, q = 1.85081 p0 = 0.96616, p1 = 0.03384, ω = 1.0 | D = 207, p < 0.0001 | M8 > M7: Evidence for positive selection if ω > 1.0 |

# Table S10. Mitogenome amino acids under selection

Amino acids inferred to be under selection in the sperm whale, numbered with respect to sperm whale reference sequence NC\_002503. The 3D representation of the proteins and 2D representation of the features/secondary structure represented in this table can be found in supplemental Figure S4. ATP8 is not modelled as a 3D protein because no homologous structures were found through SWISS-MODEL. References for 3D models: 5LDW/5LC5 (Zhu *et al.* 2016); 3ABM (Aoyama *et al.* 2009); 4D6T (Capper *et al.* 2015). The amino acid states in other cetaceans versus the sperm whale are given in the ‘Other cetaceans > sperm whale’ column.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Gene and amino acid position relative to NC\_002503** | **Base pair position relative to NC\_002503** | **Model** | **Secondary structure** | **Feature** | **Other cetaceans > sperm whale** | **Method** | **TreeSAAP Properties** |
| ND1, 102nd residue | 3,078-3,080 | 5LDW.1.H | Alpha helix | Transmembrane | A, V > I | TreeSAAP | Equilibrium constant (ionization of COOH) |
| ND2, 52nd residue | 4,096-4,098 | 5LDW.1.N | Alpha helix | -- | A > T | TreeSAAP | Alpha-helical tendencies |
| ND2, 100th residue | 4,240-4,242 | 5LDW.1.N | Alpha helix | Transmembrane | L, M, V, T > I | HyPhy | -- |
| ATP8, 30th residue | 7,902-7,904 | *NA* | *NA* | -- | K, N > M | PAML (M8), TreeSAAP | Long-range non-bonded energy, polar requirement, polarity |
| COX3, 128th residue | 9,037-9,039 | 3ABM.1.C | Turn | Directly adjacent to transmembrane | E > K | TreeSAAP | Power to be at the C-terminal |
| COX3, 184th residue | 9,205-9,207 | 3ABM.1.C | -- | Directly adjacent to transmembrane | A > T | TreeSAAP | Alpha-helical tendencies |
| COX3, 229th residue | 9,340-9,342 | 3ABM.1.C | Bend | -- | S > P | TreeSAAP | Power to be at the C-terminal |
| ND4L, 67th residue | 10,124-10,126 | 5LDW.1.K | Alpha helix | Transmembrane | A > T | TreeSAAP | Alpha-helical tendencies |
| ND4, 404th residue | 11,425-11,427 | 5LC5.1.M | Alpha helix | Transmembrane | A > T | TreeSAAP | Alpha-helical tendencies |
| ND4, 423rd residue | 11,482-11,484 | 5LC5.1.M | -- | -- | I > V | TreeSAAP | Equilibrium constant (ionization of COOH) |
| ND5, 20th residue | 11,851-11,853 | 5LDW.1.L | Alpha helix | Transmembrane | L, M, V > I | TreeSAAP | Equilibrium constant (ionization of COOH) |
| ND5, 116th residue | 12,139-12,141 | 5LDW.1.L | Alpha helix | Directly adjacent to transmembrane | Q, R > L | TreeSAAP | Buriedness, hydropathy, polarity, total non-bonded energy |
| ND5, 433rd residue | 13,090-13,092 | 5LDW.1.L | Bend | Directly adjacent to transmembrane | E, G > N | HyPhy, MEME | -- |
| ND5, 455th residue | 13,156-13,158 | 5LDW.1.L | Alpha helix | Transmembrane | K > M | TreeSAAP | Buriedness, chromatographic index, long-range non-bonded energy, mean r.m.s. fluctuation displacement, polar requirement, polarity |
| ND5, 482nd residue | 13,237-13,239 | 5LDW.1.L | -- | -- | M, T > I | TreeSAAP | Equilibrium constant (ionization of COOH) |
| ND5, 525th residue | 13,366-13,368 | 5LDW.1.L | Bend | -- | L > S | TreeSAAP | Chromatographic index, mean r.m.s. fluctuation displacement |
| ND5, 539th residue | 13,408-13,410 | 5LDW.1.L | Alpha helix | -- | A, H, L, Y > D | MEME | -- |
| ND5, 598th residue | 13,585-13,587 | 5LDW.1.L | Alpha helix | Transmembrane | A, I, S, T, V > F | MEME | -- |
| CYTB, 142nd residue | 14,622-14,624 | 4D6T.1.C | Alpha helix | Transmembrane | G > A | TreeSAAP | Alpha-helical tendencies, coil tendencies |

# Figure S1. PSMC plots and analysis methods.

1. Combined PSMC plots for different mutation rates and generation times (see methods below) for BioSample SAMN06187412 (Pacific).

../../../Mol%20Ecol%20Lab/%20Projects/%20Sperm%20whales/Pmac%20genomic%20and%20demographics/PSMC_Jan2017/PSMC_sensitivity_tests/Pmac7412_sensitivity_psmc.out.pdf

1. Bootstrapped PSMC plots for the three individual samples. Lighter colored plot lines represent 100 bootstrap plots around the PSMC plot (darker line).

Pacific (BioSample SAMN06187412)

Pmac7412r0u2.9E8_boot_x10k.out.pdf

Atlantic (BioSample SAMN01906698)

merged_6698_boot_combined.out.pdf

Indian (BioSample SAMN06187413)

merged7413_boot_combined.out_crop.pdf

1. Bootstrapped PSMC plots for three cross-ocean pseudo-diploids.

Atlantic-Pacific

mergedPmac_PD_boot_combined_10k-10M_green_crop.out.pdf

Indian-Pacific

merged_7412-7413_boot_combined.out.pdf

Indian-Atlantic

merged_6698-7413_boot_combined.out.pdf

**Methods details for genome assembly and PSMC analysis:**

The sperm whale short-read archive (SRA) data were retrieved from Genbank and filtered using Trimmomatic (Bolger *et al.* 2014) to remove poor-quality nucleotides (leading and trailing nucleotides with Q<20; sliding window of 4 nucleotides to replace bases with N if Q<15; discard reads <40bp long). Reads from each SRA file were first aligned to a reference sperm whale mitogenome (accession No. KC312603 with the first and last 40bp of the linearized sequence added to the opposite end to improve alignment of reads over the artificial break point) using BWA mem (v0.7.15-r1140; Li 2013; Li & Durbin 2009), followed by extraction of the unmapped reads from the resulting bam files using SAMTOOLS (v1.2). The resulting reads were mapped to the reference nuclear genome (accession No. GCA\_000472045.1) using BWA mem. Genome assemblies from multiple SRA files from a single BioSample were merged using SAMTOOLS. Depth of coverage for combined alignment bam files was determined using ANGSD (Korneliussen *et al.* 2014) as described by Morin et al. (2017). All four available SRA files for BioSamples SAMN06187412 and SAMN067413 were combined for respective assemblies, but only 3 of the available SRA files from BioSample SAMN01906698 were used in order to obtain approximately the same depth of coverage in all alignments.

PSMC analysis was performed on each sample after first converting the merged bam file to a diploid consensus fastq file using SAMTOOLS and BCFTOOLS (v1.2; Li 2011), and then creating a pseudo-fasta file (PSMC; Li & Durbin 2011). The default settings for humans were used for PSMC (Li & Durbin 2011), but the PSMC plot was scaled to an autosomal mutation rate (µA) of 9.1 x 10-10 substitutions per site per year (Dornburg *et al.* 2012) and a generation time of 31.9 years (Taylor *et al.* 2007), for a rate of 2.9 x 10-8 substitutions/site/generation, and we conducted 100 bootstrap resamplings.

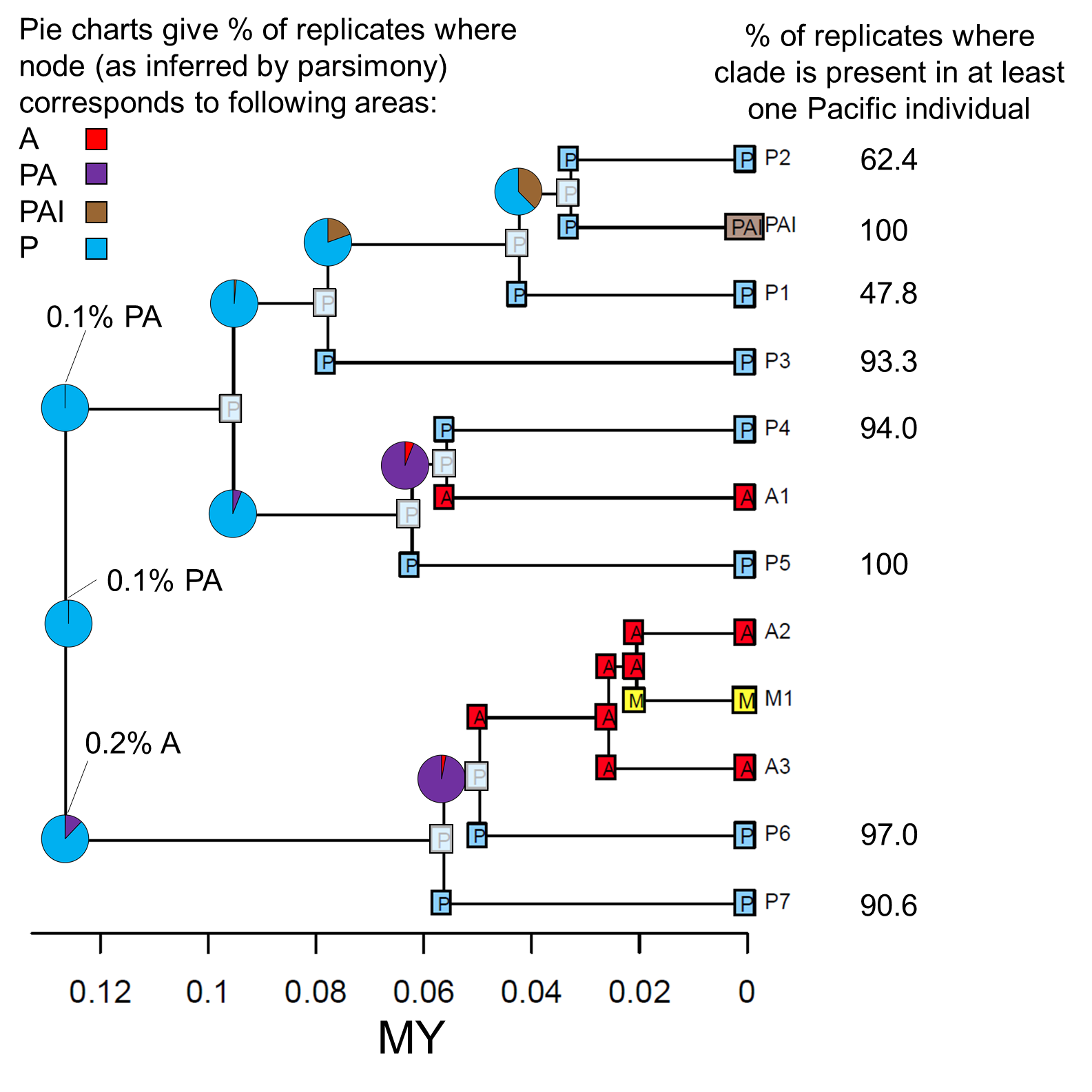
Pseudodiploids were created from combinations of two diploid consensus sequences using SEQTK mergefa. This process randomly selects alleles from heterozygous sites in each genome as the alleles in the pseudodiploid genome. PSMC was then conducted as above. PSMC bases its inference of *Ne* through time on time to most recent ancestor between alleles in independently recombining regions. There will be a uniform false detection rate (UFDR) of 0.5 across the genome of heterozygote sites that are shared in both individuals, as half the time the same allele will be selected from heterozygotes in both samples, creating a homozygote. This UFDR, which can also occur due to using low coverage sequences and mimics a lower mutation rate, typically upwardly biases recent rates, and downwardly biases more ancient coalescence. The effect will not be constant through time, due to effects of lineage sorting reducing the number of shared heterozygotes. This likely explains the imperfect estimation of ancestral *Ne* between 0.15-1Myr

Sensitivity testing of PSMC was conducted using three mutation rates and two generation times. The generation times were estimated by Taylor et al. (2007), based on a stable population (31.9 yr) and a growing population (26.5 yr). The mutation rates included the median and 95% CI rates for odontocetes (Dornburg *et al.* 2012) scaled by the sperm whale generation time:

|  |  |  |
| --- | --- | --- |
| per year | per generation (31.9) | per generation (26.5) |
| 9.10E-10 | 2.90E-08 | 2.41E-08 |
| 6.68E-10 | 2.13E-08 | 1.77E-08 |
| 1.18E-09 | 3.76E-08 | 3.13E-08 |

# Figure S2: Subsampling BioGeoBears clades.

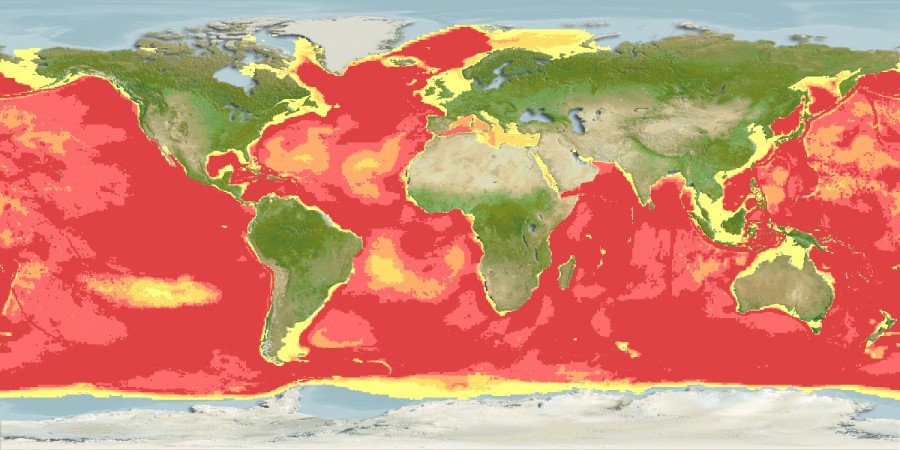
The impact of down-sampling individuals from the Pacific on ancestral area reconstruction. Using subsampling\_clades v.1.0.0 (<https://github.com/laninsky/subsampling_clades>), we examined the clades that were present in the Pacific after down-sampling to 35 individuals (the same sample size as for the Atlantic, excluding the Gulf of Mexico and Mediterranean) for *n* = 1,000 permutations. The proportion of permutations where each clade was present for the Pacific is given to the right of the tree (taken from the summary\_record output from subsampling\_clades). The pie graphs on the nodes of the trees give the ancestry area reconstruction based on parsimony across the 1,000 replicates (inferred from the full\_clade\_record output from subsampling\_clades). Inferences of ancestral node areas that differ between the full dataset and the subsampled datasets are shown as pale squares. Only 1 out of the 1,000 down-sampled Pacific replicates (0.1%) failed to infer the Pacific as the location of the root of the tree, indicating that the inference of this location for the root is not due only to uneven sampling.



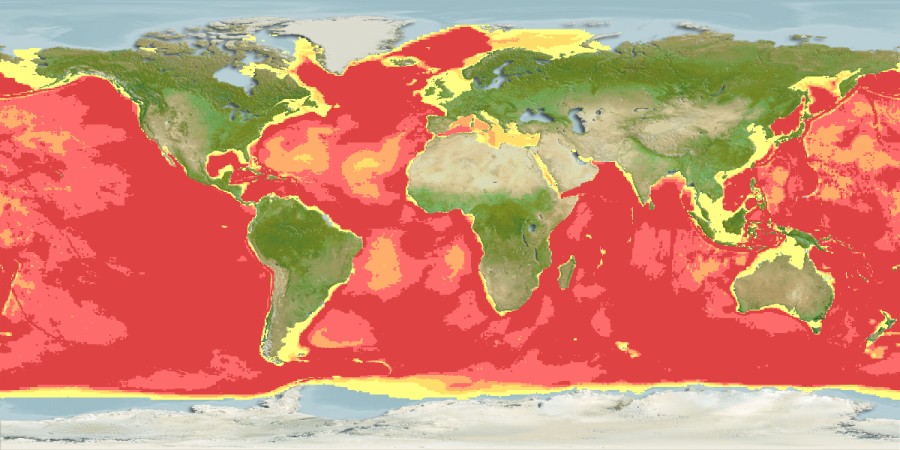
# Figure S3. AquaMaps habitat models

Habitat suitability plots for male and female sperm whales are shown separately, based on differential use of colder high-latitude regions. Last glacial maximum (LGM) and future (year 2100) models do not include the primary production parameter envelope as data are not available for those models. The male distribution is based on the global habitat preferences of sperm whales regardless of sex (Kaschner *et al.* 2016). The female model was altered by changing the minimum and preferred minimum sea surface temperatures (SST) to 5°C and 12°C, respectively, to limit the female range to latitudes that approximate their maximum summer SST range, as inferred from commercial whaling data (Ivashchenko *et al.* 2014). Habitat not inferred to be suitable shown in blue, less preferred in yellow through to most prepared in red.

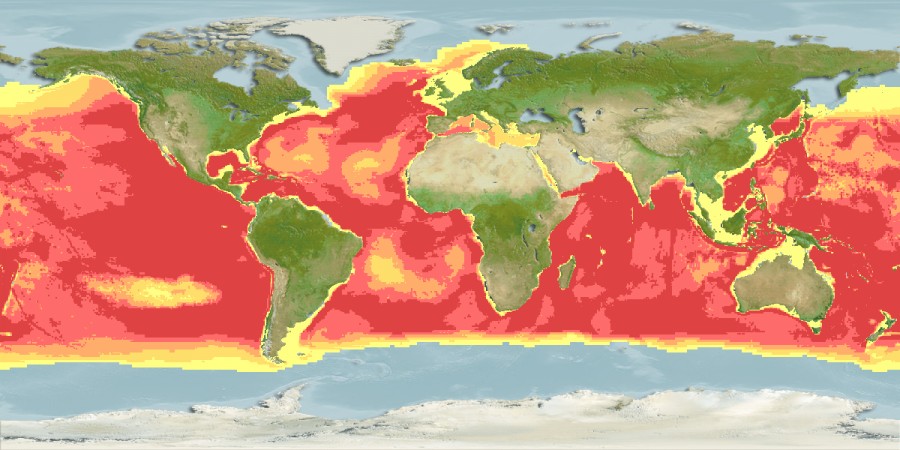
A) Present day male habitat model with primary production parameter envelope.



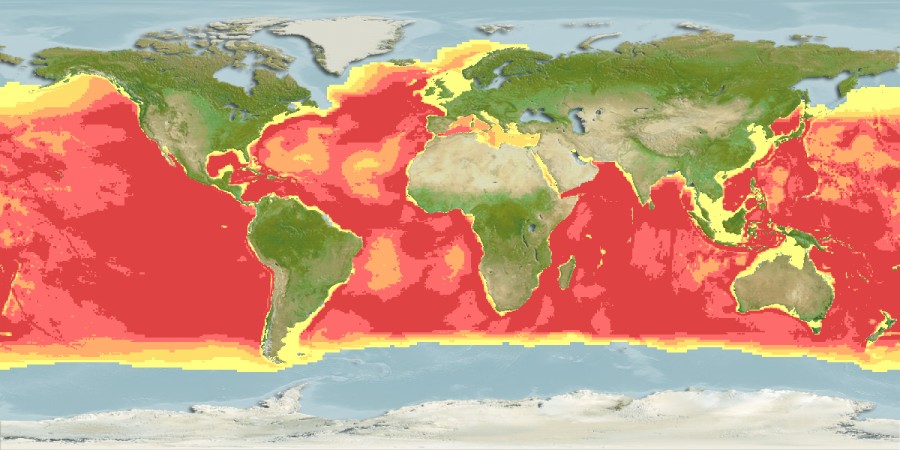
B) Present day male habitat model without primary production parameter envelope.



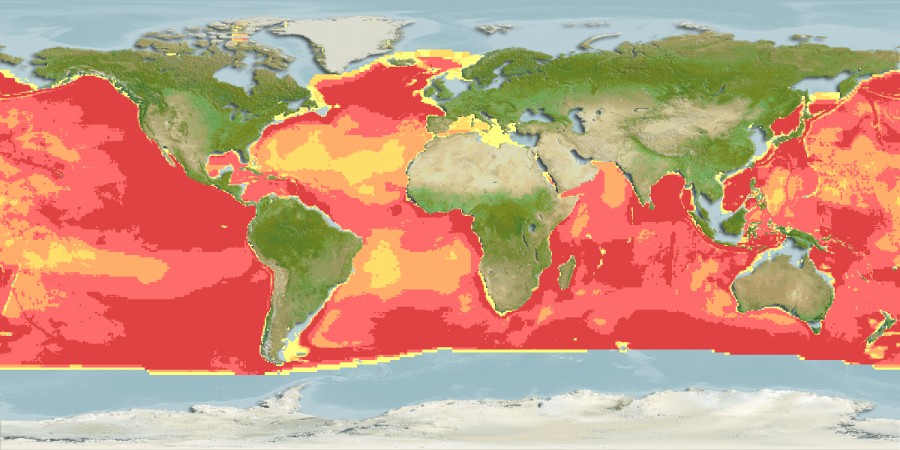
C) Present day female habitat model with primary production parameter envelope.



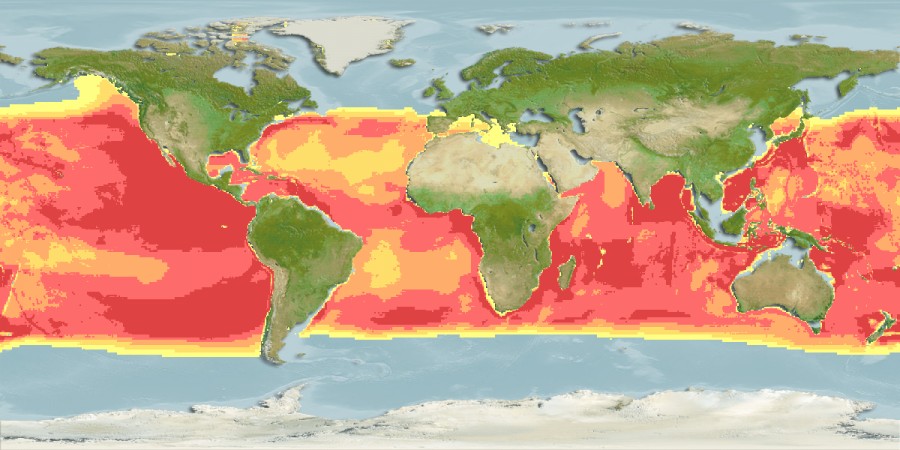
D) Present day female habitat model without primary production parameter envelope.

****

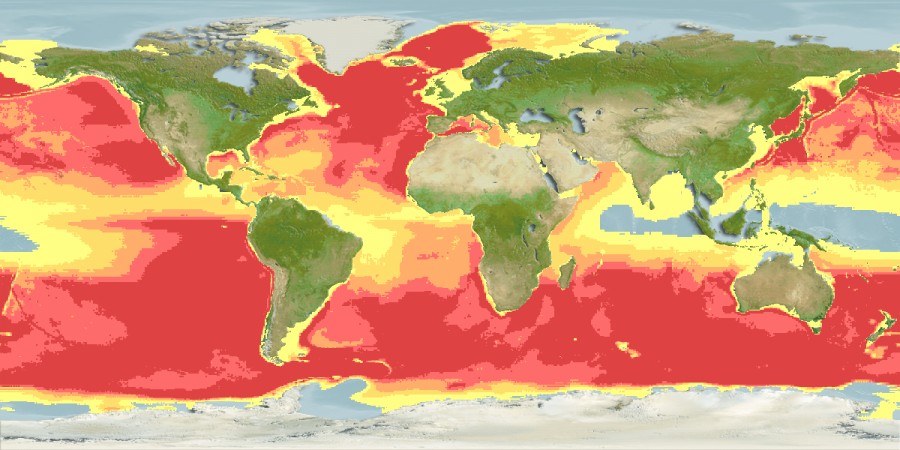
E) LGM male habitat model.

****

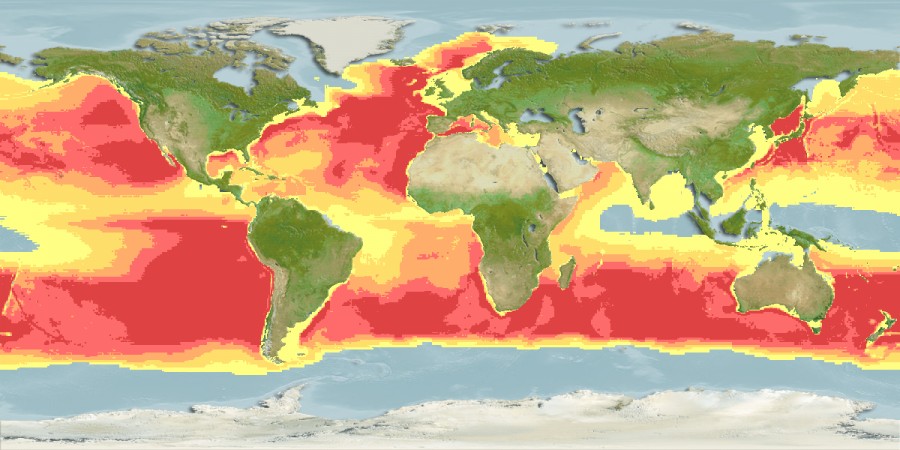
F) LGM female habitat model.

****

G) Year 2100 male habitat model.

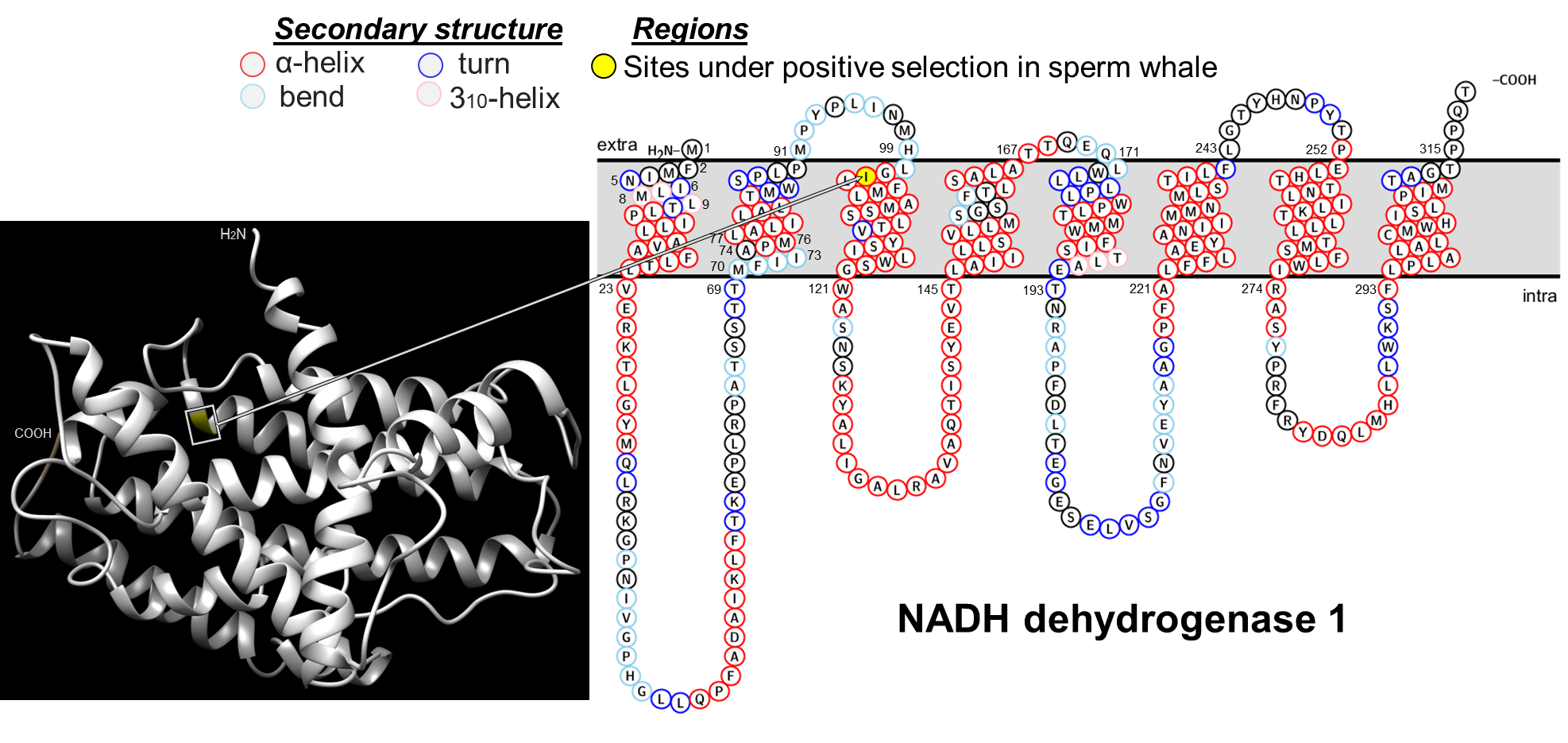
****

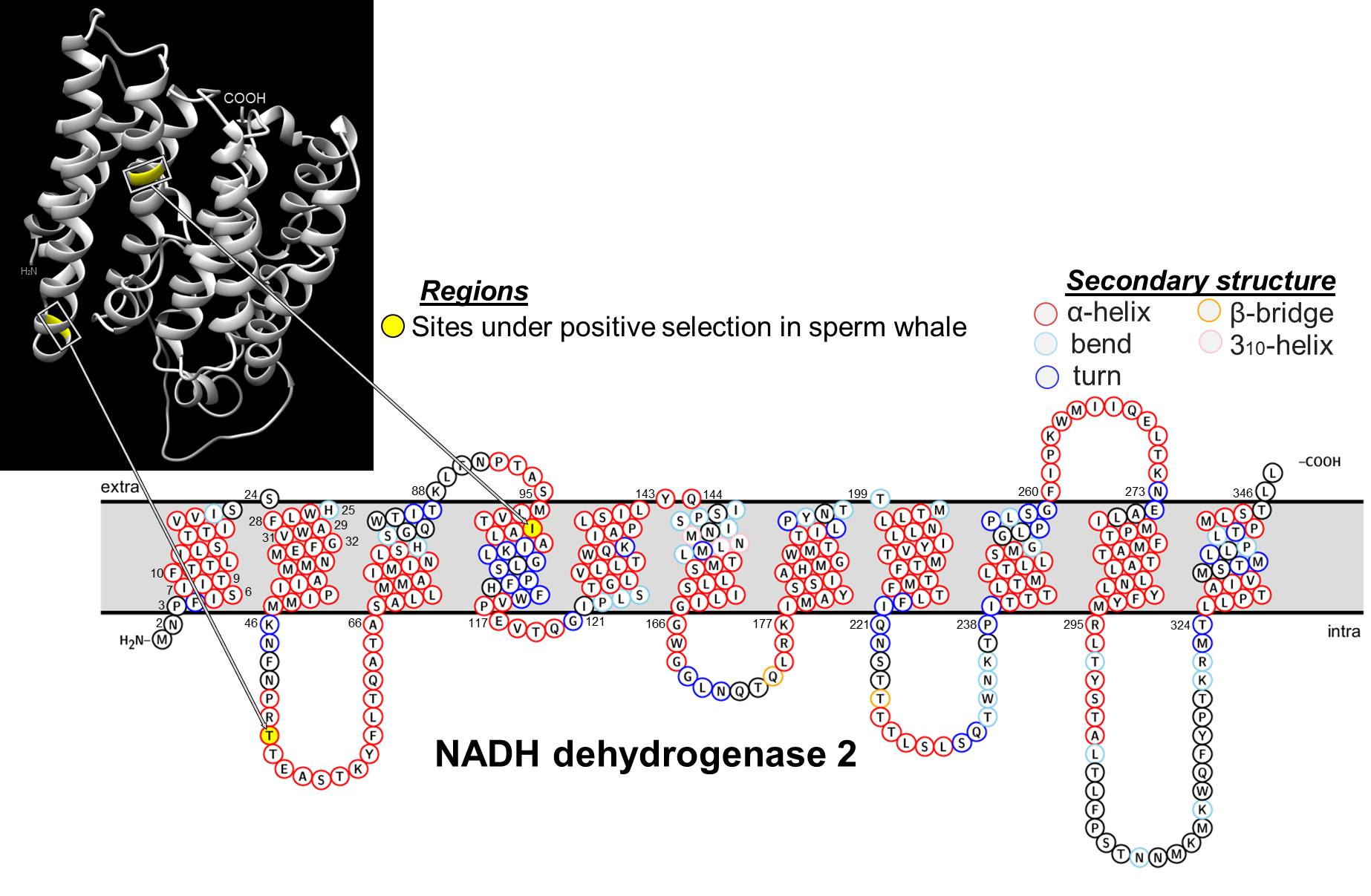
H) Year 2100 female habitat model.

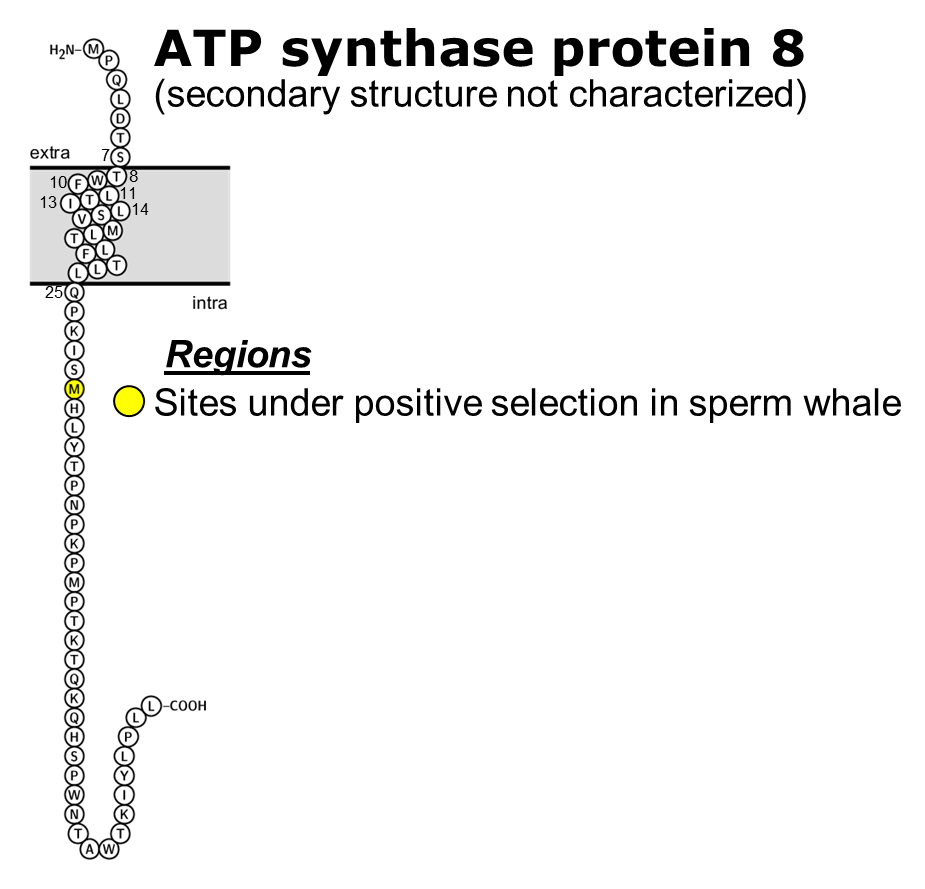
****

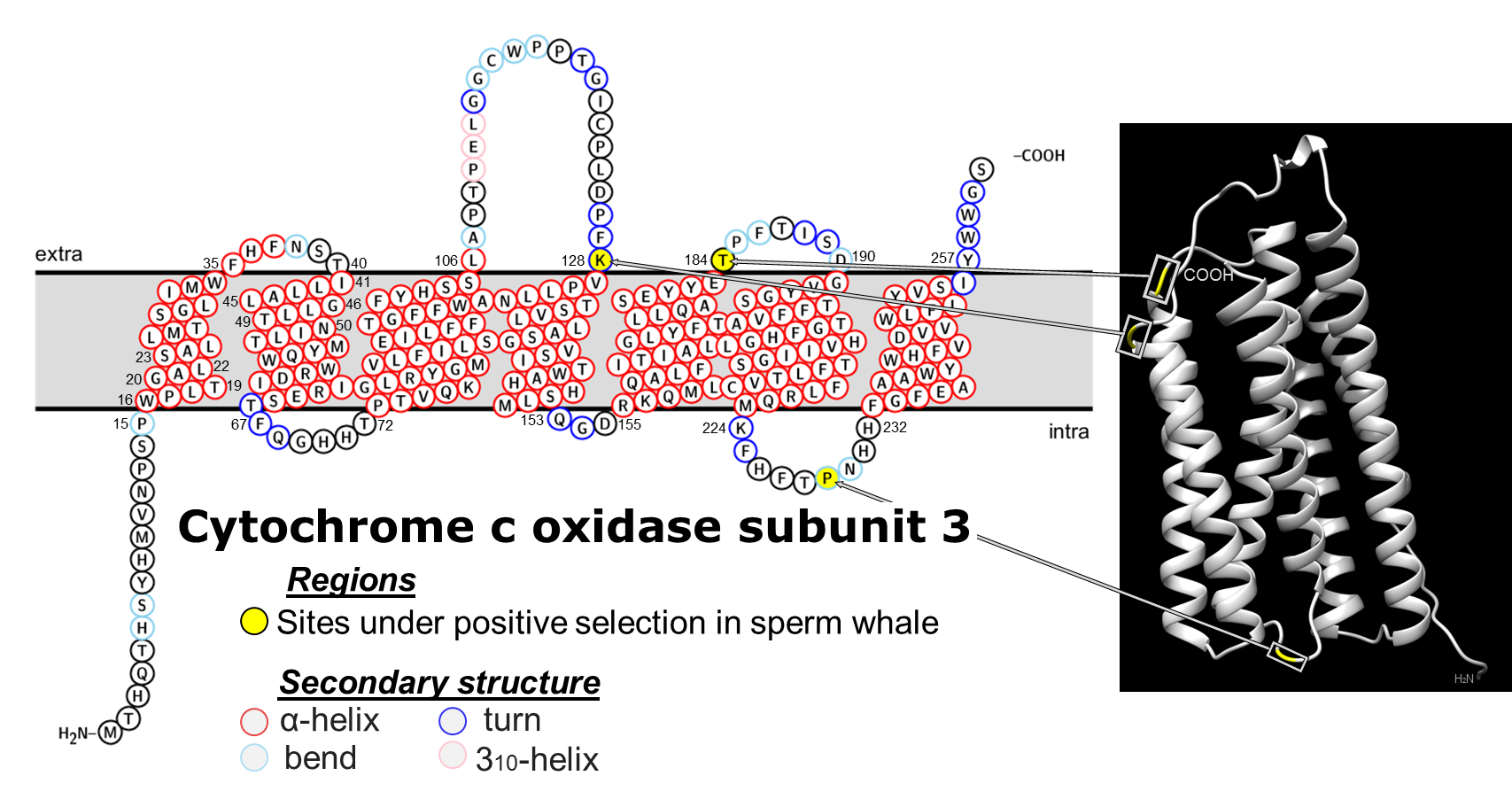
# Figure S4. Sites under positive selection

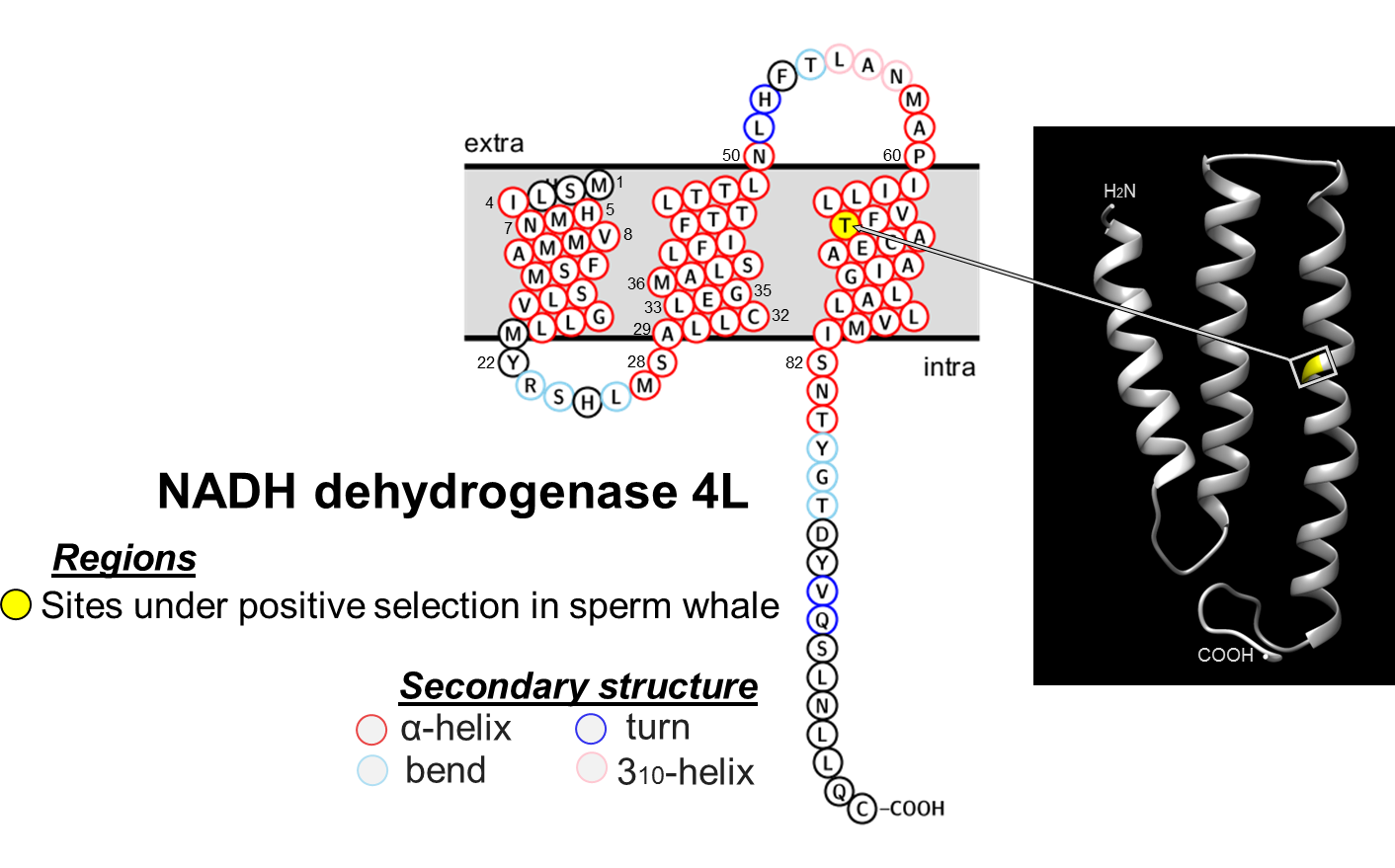
Visualization of the proteins that have amino acids inferred to be under selection in the sperm whale. The amino acid sequence for each protein is taken from the Refseq sperm whale sequence NC\_002503. The 3D models used to annotate the sperm whale proteins are listed in Table S10. For each protein, residues directly adjacent to transmembrane regions are numbered. Transmembrane residue numbering follows the same pattern as the first ascending and descending transmembrane regions which are partially numbered for each protein. No 3D model available for ATP8.

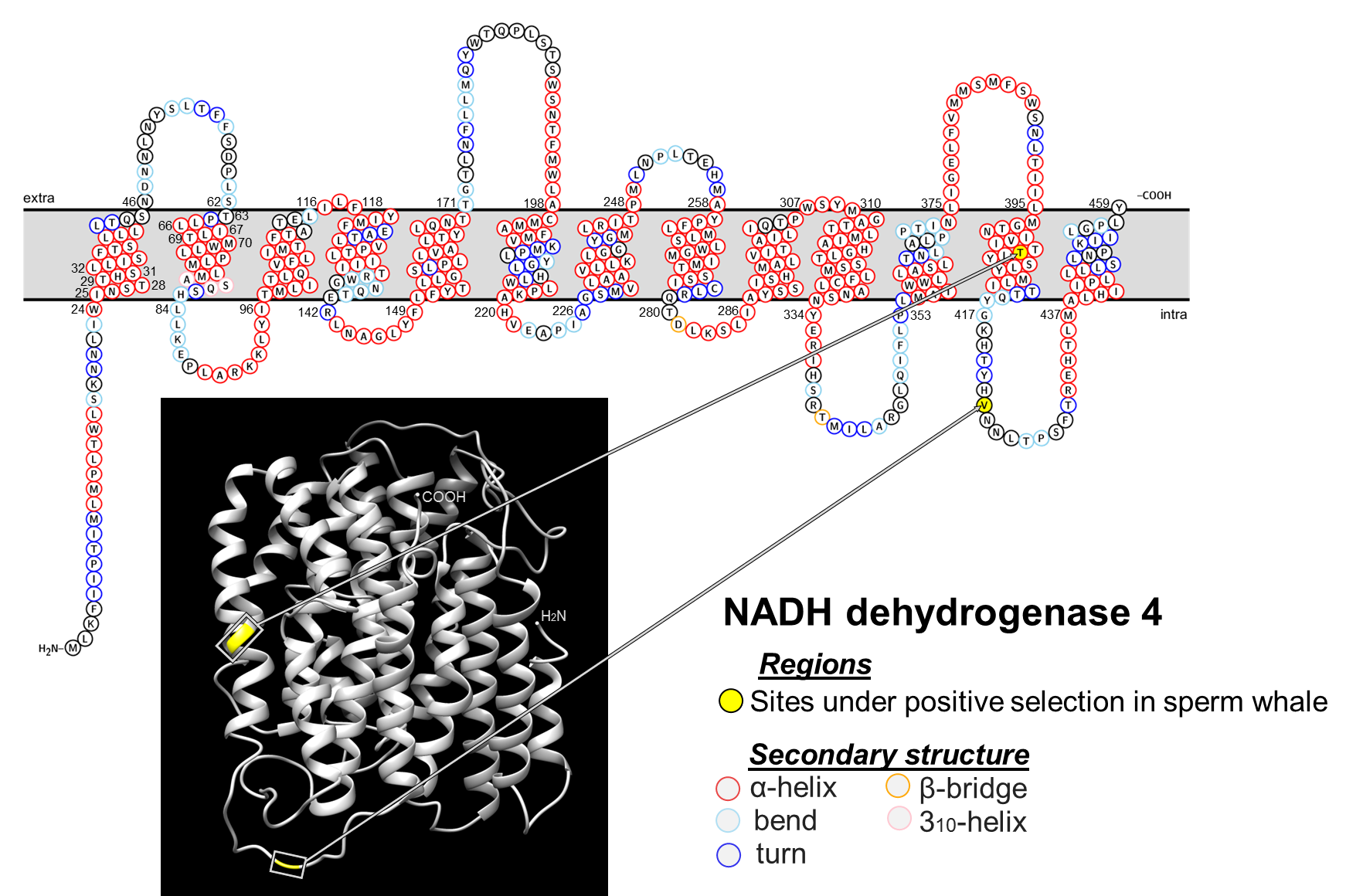


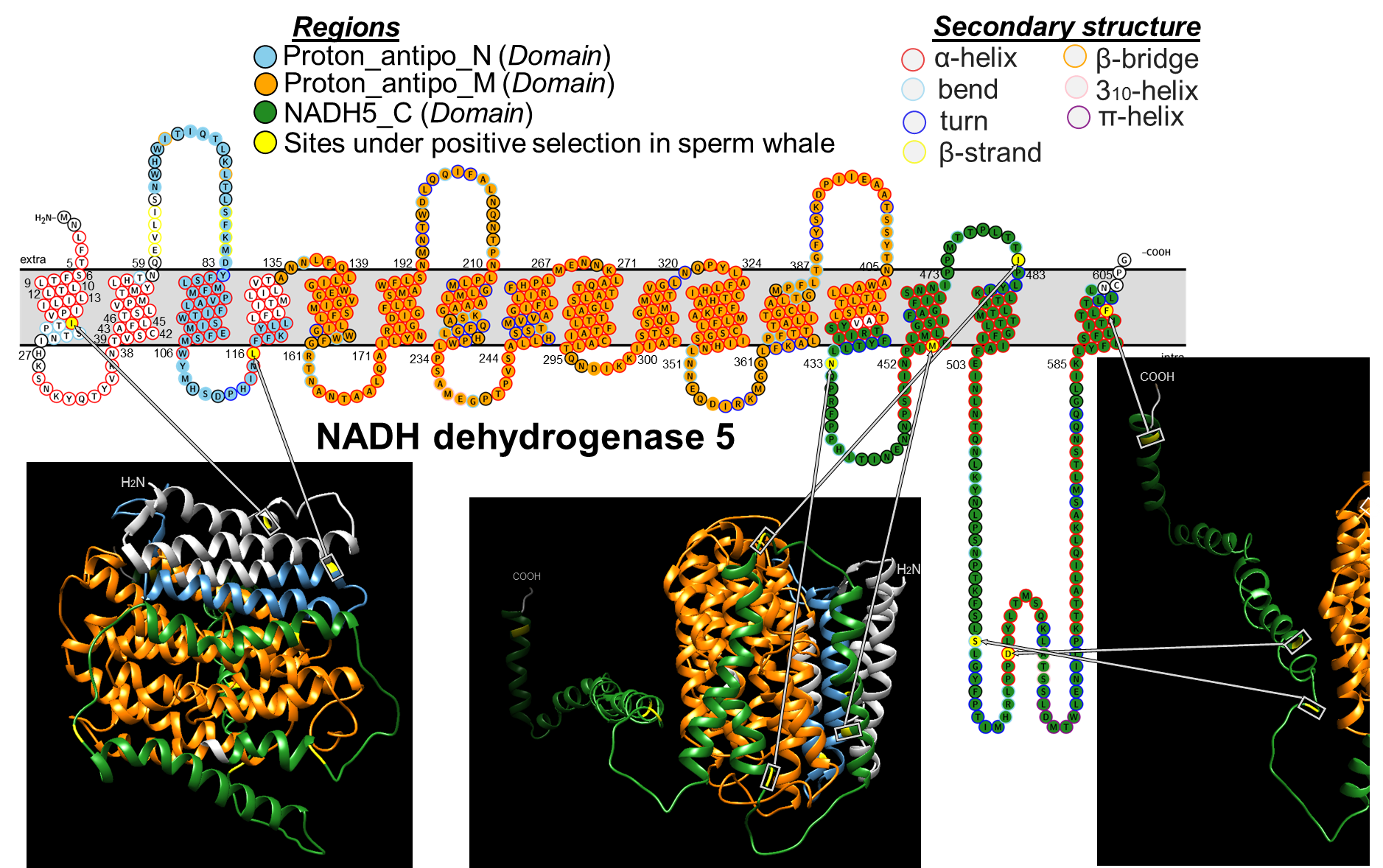


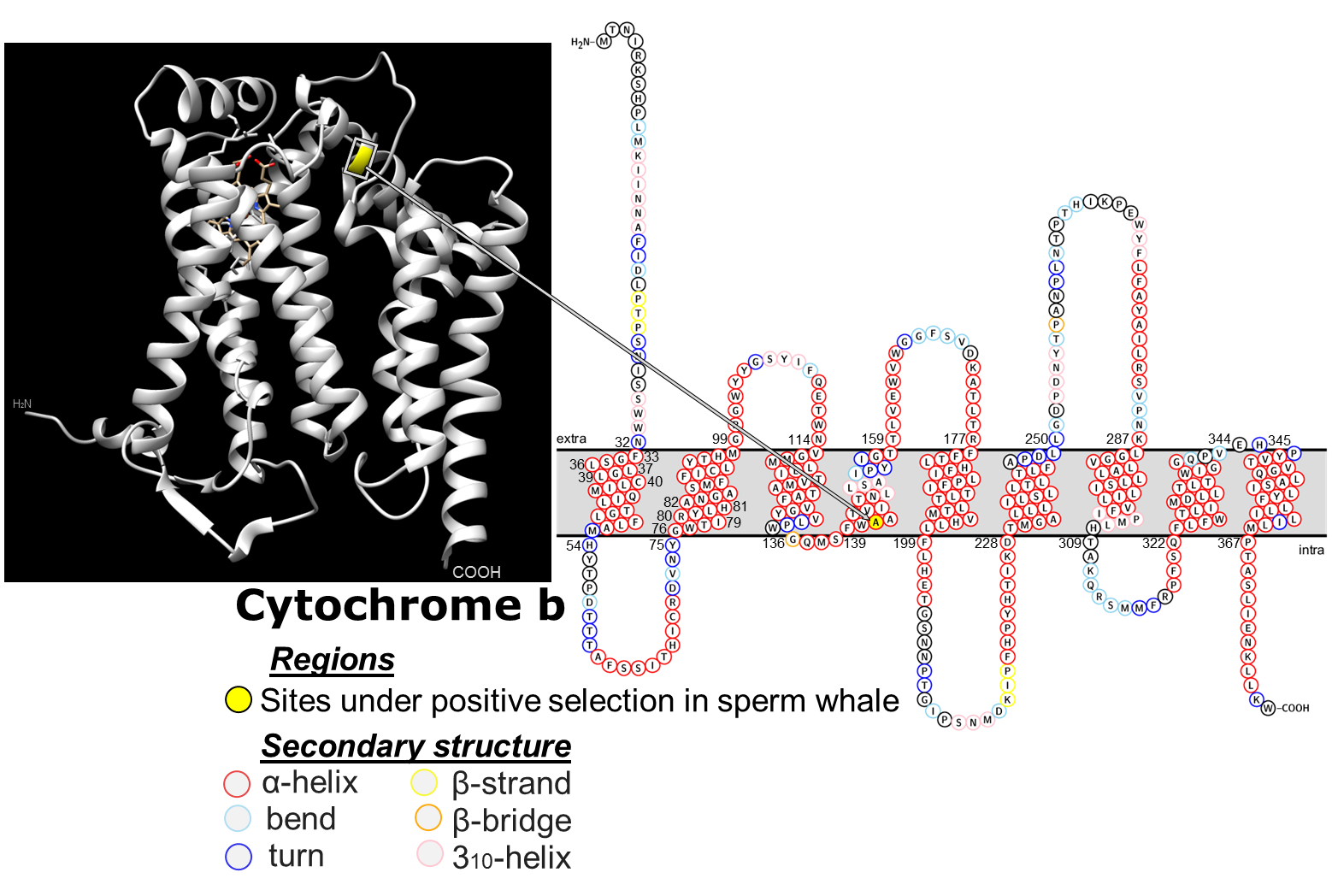












# Figure S5. BEAST Phase I phylogenetic tree.

Time calibrated phylogenetic tree of sperm whales based on divergence of Ziphiidae and Odontoceti. Nodes/scale bar show the divergence time estimates (in millions of years) and the 95% HPDI. Posterior support of all nodes (except within the sperm whale clade, where individual haplotype ID’s are indicated) were ≥0.99.

Macintosh HD:Users:phil.morin:Documents:Mol Ecol Lab: Projects: Sperm whales:Pmac_mitogenomics:Pmac_wOutgroups_uniqueHaps_beast_010715:Phase1_beast_w2_outgroup:4Pmac_wZiphiidae_Kbrev_200715_concat_modsites_100M.tre.pdf

# Figure S6. Skyline plots for ocean basins.

Skyline plots based on only samples from within the A) Atlantic (N=50) and B) Pacific (N=124) Oceans. All other parameters were as described in the methods for skyline analysis of the full mitogenome data set. Median TMRCA was 86 kyr (95% CI 57 – 119 kyr) for the Pacific, and 78 kyr (49 – 112 kyr) for the Atlantic.

A)

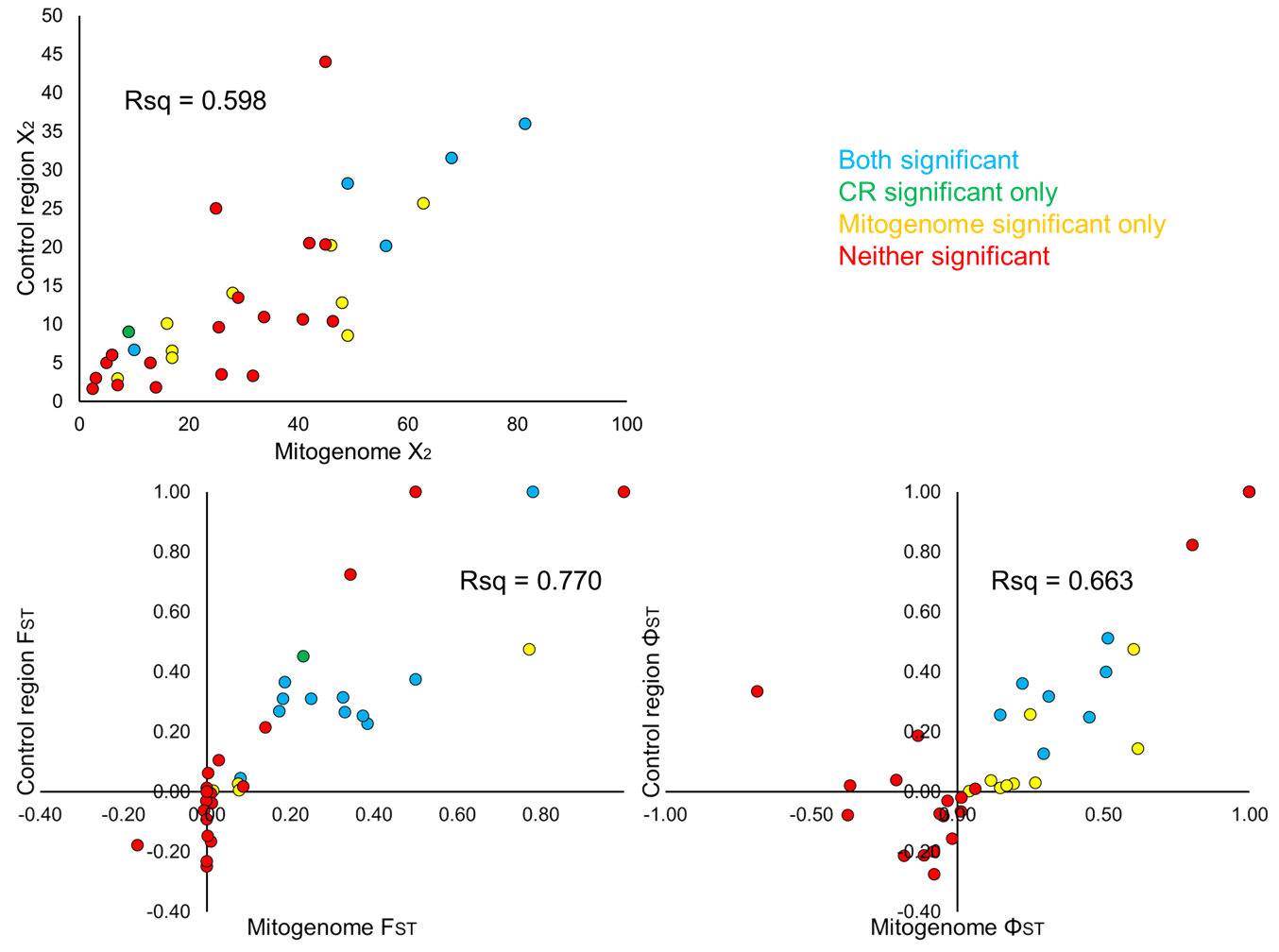
../../../Mol%20Ecol%20Lab/%20Projects/%20Sperm%20whales/Pmac_mitogenomics/Skyride_analyses_Jan2016/Alana_skyline-skyride_250117/Atl_pma_mito_27Feb2017_skyline_plot.pdf

B)

../../../Mol%20Ecol%20Lab/%20Projects/%20Sperm%20whales/Pmac_mitogenomics/Skyride_analyses_Jan2016/Alana_skyline-skyride_250117/Pac_pma_mito_27Feb2017_skyline_plot.pdf

# Figure S7. Mitogenome vs. control region population structure.

Pairwise divergence for mitochondrial genome haplotypes vs. 619bp control region haplotypes for regions indicated in Table S6. Plots show divergence values color coded by whether each pairwise comparison was significant (p < 0.05) for each metric. Pairwise divergence was calculated using the strataG package in R (Archer *et al.* 2017). Although more comparisons were significant based on mitogenomes than control regions for all three metrics, only χ2 had significantly more significant tests (based on a paired T-test).



# Supplemental methods

R-script for haplotypic diversity with bootstrap

###########

#devtools::install\_github('ericarcher/strataG')

# calculate haplotype diversity and nucleotide divergence within populations

# (usually called "nucleotide diversity"), and bootstrap results for significant

# differences in values between strata.

# Requires fasta file of unique haplotype sequences and .csv file of strata schemes,

# including sample ID's ("id"), stratificaiton schemes, and haplotype ID's ("Mitogenome\_Hap").

# Requires strataG 2.0.0

###########

rm(list = ls())

library(strataG)

library(parallel)

Pmac\_haps <- read.fasta("80\_Pmac\_Unique\_hap\_sequences\_fasta\_010715.fasta")

strata.df <- readGenData("AllPmac\_strata.csv")

rownames(strata.df) <- strata.df$id

strata.df$ap <- strata.df$global\_1HapPerGroup

strata.df$ap[strata.df$ocean2 %in% c("GoM", "Mediterranean")] <- "Atlantic"

st <- "global\_1HapPerGroup"

nrep <- 1000

num.cores = 2

g.df <- na.omit(strata.df[, c("id", st, "Mitogenome\_Hap")])

g <- df2gtypes(g.df, ploidy = 1, sequences = Pmac\_haps)

HapDiv <- exptdHet(g)

HapDiv2 <- data.frame(KeyName=names(HapDiv), value=HapDiv, row.names=NULL)

df <- sapply(HapDiv2, as.numeric)#df <- rownames(df) <- rownames(HapDiv2)

write.csv(df, paste(st,"\_HapDiv.csv",sep = ""))

#nucleotide diversity (should be calculated on the whole population set of sequences,

#not unique haplotypes; see note below from Eric(3/4/16) that describes function)

model <- "TN93" #specify optimal model from jmodeltest results.

gamma <- 1.5 #

nd <- nucleotideDivergence(g, model = model, gamma = gamma, variance = FALSE,

pairwise.deletion = TRUE)

nd2 <- data.frame(KeyName=names(nd), value=nd)

## can't get this to write to csv, so just copy value from nd2 Object.

#nd2 <- sapply(nd2, as.numeric)

#nd2 <- as.matrix(nd2)

#nd2 <- sapply(nd2,unlist)

#write.csv(nd2$Mitogenome\_Hap$within, paste(description,"\_",st,"\_",model,"\_NucDiv\_w-in.csv",

sep=""))

#write.csv(nd2$Mitogenome\_Hap$between, paste(description,"\_",st,"\_",model,"\_NucDiv\_dA.csv",

sep=""))

#####################

boot.dist <- mclapply(1:nrep, function(i) {

boot.i <- unname(unlist(tapply(1:nrow(g.df), g.df[, st], function(x) sample(x, rep = T))))

boot.df <- g.df[boot.i, c("id", st, "Mitogenome\_Hap")]

boot.df$id <- 1:nrow(boot.df)

boot.g <- df2gtypes(boot.df, ploidy = 1, sequences = Pmac\_haps)

het <- sapply(strataSplit(boot.g), exptdHet)

names(het) <- gsub(".Mitogenome\_Hap", "", names(het))

nuc.div <- nucleotideDivergence(

boot.g, model = "TN93", gamma = 1.5, pairwise.deletion = TRUE

)[[1]]$within[, "mean"]

cbind(het = het, nuc.div = nuc.div)

}, mc.cores = num.cores)

boot.het <- t(sapply(boot.dist, function(x) x[, "het"]))

boot.nuc.div <- t(sapply(boot.dist, function(x) x[, "nuc.div"]))

obs.het <- sapply(strataSplit(g), exptdHet)

names(obs.het) <- gsub(".Mitogenome\_Hap", "", names(obs.het))

obs.nuc.div <- nucleotideDivergence(

g, model = "TN93", gamma = 1.5, pairwise.deletion = TRUE

)[[1]]$within[, "mean"]

st.pairs <- data.frame(t(combn(strataNames(g), 2)), stringsAsFactors = F)

colnames(st.pairs) <- c("strata.1", "strata.2")

n <- c(table(strata(g)))

st.pairs <- cbind(st.pairs, n.1 = n[st.pairs$strata.1], n.2 = n[st.pairs$strata.2])

het.df <- st.pairs

het.df <- cbind(het.df, het.1 = obs.het[het.df$strata.1], het.2 = obs.het[het.df$strata.2])

het.df$obs.het.diff <- -apply(het.df, 1, function(x) diff(obs.het[x[1:2]]))

het.df$pct.gte.0 <- apply(het.df, 1, function(x) {

boot.diff <- -apply(boot.het[, x[1:2]], 1, diff)

(sum(boot.diff >= 0) + 1) / (length(boot.diff) + 1)

})

nuc.div.df <- st.pairs

nuc.div.df <- cbind(

nuc.div.df,

nuc.div.1 = obs.nuc.div[nuc.div.df$strata.1],

nuc.div.2 = obs.nuc.div[nuc.div.df$strata.2]

)

nuc.div.df$obs.nuc.div.diff <- -apply(nuc.div.df, 1, function(x) diff(obs.nuc.div[x[1:2]]))

nuc.div.df$pct.gte.0 <- apply(nuc.div.df, 1, function(x) {

boot.diff <- -apply(boot.nuc.div[, x[1:2]], 1, diff)

(sum(boot.diff >= 0) + 1) / (length(boot.diff) + 1)

})

rownames(het.df) <- rownames(nuc.div.df) <- NULL

save.image(paste0("bootstrap ws\_", st, ".rdata"))

write.csv(

het.df,

file = paste0("haplotypic diversity bootstrap results\_", st, ".csv"),

row.names = F

)

write.csv(

nuc.div.df,

file = paste0("nucleotide divergence bootstrap results\_", st, ".csv"),

row.names = F

)

############

# Supplemental Materials References

Alexander A, Steel D, Hoekzema K, Mesnick SL, Engelhaupt D, Kerr I, . . . Baker CS (2016) What influences the worldwide genetic structure of sperm whales (*Physeter macrocephalus*)? *Molecular Ecology* **25**, 2754-2772.

Alexander A, Steel D, Slikas B, Hoekzema K, Carraher C, Parks M, . . . Baker CS (2013) Low diversity in the mitogenome of sperm whales revealed by next-generation sequencing. *Genome Biology and Evolution* **5**, 113-129.

Aoyama H, Muramoto K, Shinzawa-Itoh K, Hirata K, Yamashita E, Tsukihara T, . . . Yoshikawa S (2009) A peroxide bridge between Fe and Cu ions in the O2 reduction site of fully oxidized cytochrome c oxidase could suppress the proton pump. *Proc Natl Acad Sci U S A* **106**, 2165-2169.

Archer FI, Adams PE, Schneiders BB (2017) strataG: An R package for manipulating, summarizing, and analyzing population genetic data. *Molecular Ecology Resources* **17**, 5-11.

Arnason U, Gullberg A (1993) Comparison between the complete mtDNA sequences of the blue and the fin whale, two species that can hybridize in nature. *J Mol Evol* **37**, 312-322.

Arnason U, Gullberg A, Janke A (2004) Mitogenomic analyses provide new insights into cetacean origin and evolution. *Gene* **333**, 27-34.

Arnason U, Gullberg A, Widegren B (1991) The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Balaenoptera physalus. *J Mol Evol* **33**, 556-568.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Capper MJ, O'Neill PM, Fisher N, Strange RW, Moss D, Ward SA, . . . Antonyuk SV (2015) Antimalarial 4(1H)-pyridones bind to the Qi site of cytochrome bc1. *Proc Natl Acad Sci U S A* **112**, 755-760.

Dornburg A, Brandley MC, McGowen MR, Near TJ (2012) Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). *Molecular Biology and Evolution* **29**, 721-736.

Foote AD, Vilstrup JT, De Stephanis R, Verborgh P, Abel Nielsen SC, Deaville R, . . . Piertney SB (2011) Genetic differentiation among North Atlantic killer whale populations. *Mol Ecol* **20**, 629-641.

Hassanin A, Delsuc F, Ropiquet A, Hammer C, Jansen van Vuuren B, Matthee C, . . . Couloux A (2012) Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *C R Biol* **335**, 32-50.

Ivashchenko YV, Brownell RL, Clapham PJ (2014) Distribution of Soviet catches of sperm whales Physeter macrocephalus in the North Pacific. *Endangered Species Research* **25**, 249-263.

Kaschner K, Kesner-Reyes K, Garilao C, Rius-Barile J, Rees T, Froese R (2016) AquaMaps: Predicted range maps for aquatic species. *World wide web electronic publication,* [http://www.aquamaps.org/](http://www.aquamaps.org)*, Version 08/2016.*

Kaschner K, Quick NJ, Jewell R, Williams R, Harris CM (2012) Global coverage of cetacean line-transect surveys: status quo, data gaps and future challenges. *PLoS ONE* **7**, e44075.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356.

Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993.

Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. In: *arXiv:1303.3997v1 [q-bio.GN]*.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493-496.

Liu X, Chen R, Li G, Li J (2016) Mitochondrial genome of the Neophocaena asiaeorientalis asiaeorientalis (Phocaenidae: Neophocaena). *Mitochondrial DNA A DNA Mapp Seq Anal* **27**, 3145-3146.

Mesnick S, Taylor B, Archer EI, Martien K, Escorza Treviño S, Hancock-Hanser BL, . . . Morin PA (2011) Sperm whale population structure in the eastern North Pacific inferred by the use of single nucleotide polymorphisms (SNPs), microsatellites and mitochondrial DNA. *Molecular Ecology Resources* **11 (suppl. 1)**, 278-298.

Morin PA, Archer FI, Foote AD, Vilstrup J, Allen EE, Wade P, . . . Harkins T (2010) Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. *Genome Research* **20**, 908-916.

Morin PA, Duchene S, Lee N, Durban J, Claridge D (2012) Preliminary analysis of mitochondrial genome phylogeography of Blainville’s, Cuvier’s and Gervais’ beaked whales, p. 17. International Whaling Commission, International Whaling Commission, Scientific Meeting 64, Panama City, Panama.

Morin PA, Foote AD, Hill CM, Simon-Bouhet B, Lang AR, Louise M (2017) SNP discovery from single and multiplex genome assemblies of non-model organisms. In: *Methods in Molecular Biology: Next-Generation Sequencing* (ed. Head SR). Springer.

Moura AE, Nielsen SC, Vilstrup JT, Moreno-Mayar JV, Gilbert MT, Gray HW, . . . Hoelzel AR (2013) Recent diversification of a marine genus (Tursiops spp.) tracks habitat preference and environmental change. *Systematic Biology* **62**, 865-877.

Sasaki T, Nikaido M, Hamilton H, Goto M, Kato H, Kanda N, . . . Okada N (2005) Mitochondrial phylogenetics and evolution of mysticete whales. *Systematic Biology* **54**, 77-90.

Sasaki T, Nikaido M, Wada S, Yamada TK, Cao Y, Hasegawa M, Okada N (2006) Balaenoptera omurai is a newly discovered baleen whale that represents an ancient evolutionary lineage. *Mol Phylogenet Evol* **41**, 40-52.

Taylor BL, Chivers SJ, Larese J, Perrin WF (2007) Generation length and percent mature estimates for IUCN assessments of cetaceans. Administrative Report, Southwest Fisheries Science Center, 8604 La Jolla Shores Blvd., La Jolla, CA 92038, USA.

Thompson KF, Patel S, Williams L, Tsai P, Constantine R, Baker CS, Millar CD (2016) High coverage of the complete mitochondrial genome of the rare Gray's beaked whale (Mesoplodon grayi) using Illumina next generation sequencing. *Mitochondrial DNA A DNA Mapp Seq Anal* **27**, 128-129.

Vilstrup JT, Ho SYW, Foote AD, Morin PA, Kreb D, Krützen M, . . . Gilbert MTP (2011) Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae. *BMC Evolutionary Biology* **11**, 65.

Xiong Y, Brandley MC, Xu S, Zhou K, Yang G (2009) Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. *BMC Evol Biol* **9**, 20.

Xu TJ, Zhang B, Wang RX (2013) Direct Genbank submission.

Yan J, Zhou K, Yang G (2005) Molecular phylogenetics of 'river dolphins' and the baiji mitochondrial genome. *Mol Phylogenet Evol* **37**, 743-750.

Yao CJ, Chen CH, Hsiao CD (2016) The complete mitogenome of Ginkgo-toothed beaked whale (Mesoplodon ginkgodens) (Chordata: Ziphiidae). *Mitochondrial DNA A DNA Mapp Seq Anal* **27**, 2846-2847.

Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. *Nature* **536**, 354-358.