

Title:

Medaka population genome structure and demographic history described via genotyping-by-sequencing

Authors and Affiliations:

Takafumi Katsumura^{1,2*}, Shoji Oda³, Mitani Hiroshi³, Hiroki Oota^{1*}.

¹ Department of Anatomy, Kitasato University School of Medicine

1-15-1 Kitasato, Minami-Ku Sagamihara, Kanagawa 252-0374, Japan

² Graduate School of Natural Science and Technology, Okayama University, 1-1-1

Tsushima-naka, Kita-Ku, Okayama 700-8530, Japan

³ Department of Integrated Biosciences, Graduate School of Frontier Sciences, University of

Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan

***Authors for Correspondence:**

Takafumi Katsumura, Graduate School of Natural Science and Technology, Okayama

University, Okayama, Japan, +81 86 252 7860, tk@okayama-u.ac.jp

Hiroki Oota, Department of Anatomy, Kitasato University School of Medicine, Kanagawa,

Japan, +81 42 778 9022, hiroki_oota@med.kitasato-u.ac.jp

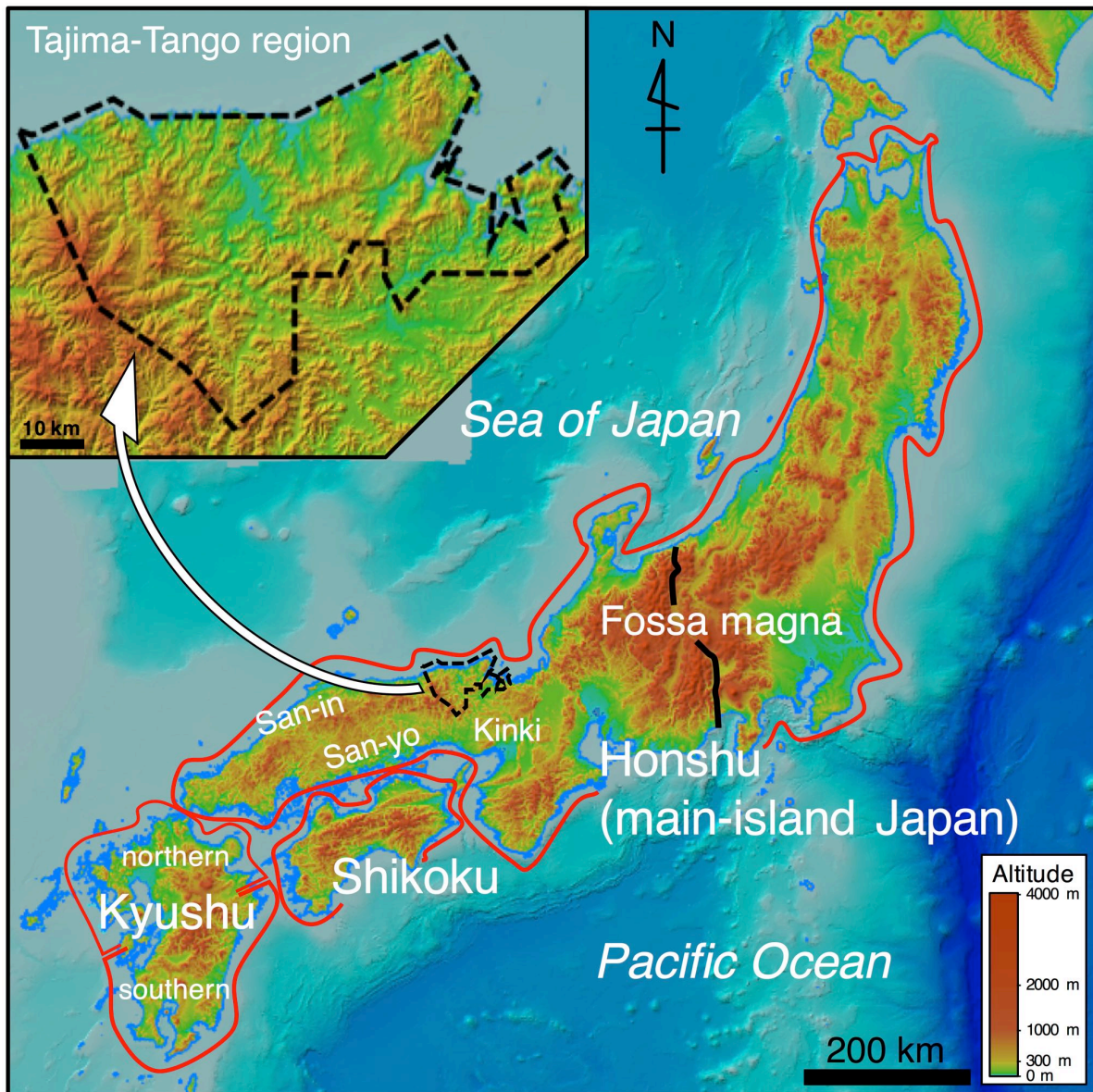
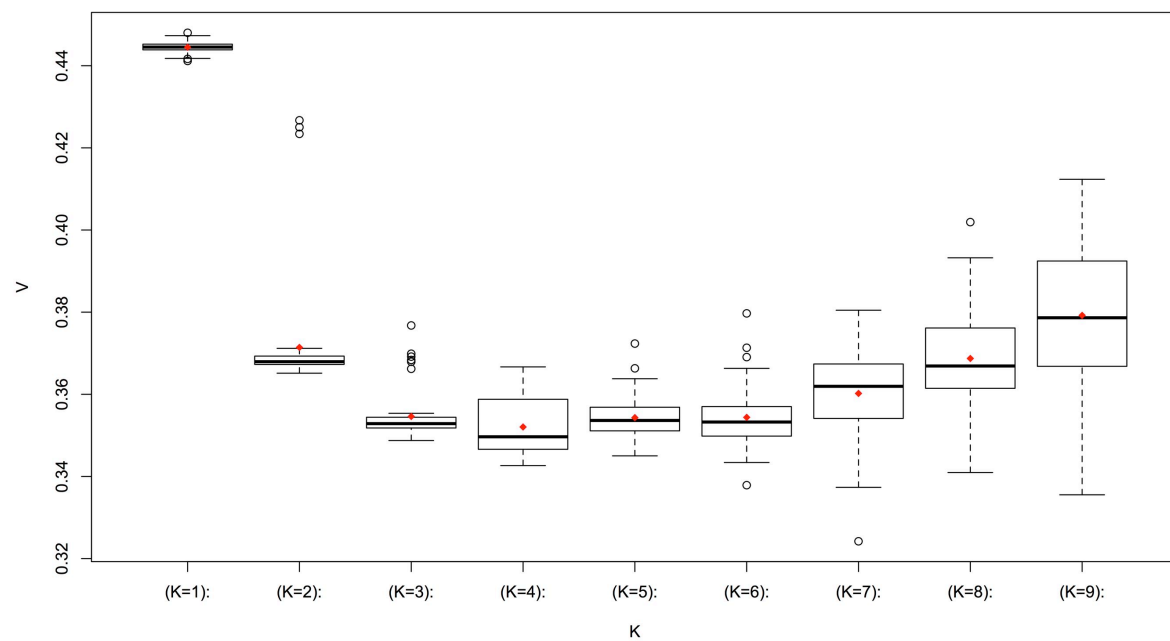


Figure S1: Geographical map of Japan. The map from the Geospatial Information Authority (GSI) of Japan (<https://maps.gsi.go.jp/development/ichiran.html>) is shown. Red lines show three islands—Honshu (main island), Shikoku and Kyushu—in the Japanese archipelago, which are known as Hondo, (mainland) Japan. In addition, regional names are shown on the west part of Honshu: Kinki, San-yo and San-in. San-yo and San-in are separated by the mountain range named Chugoku Sanchi. San-yo on the Seto Inland Sea side is mild in comparison with the San-in on the side of the Sea of Japan, and little snow falls on the coastal area. Black solid and dashed lines indicate the Fossa magna and Tajima-Tango regions, respectively.

32
33
34
35
36
37



38
39
40
41

Figure S2: Boxplot of cross-validation errors. The red point in the box represents the mean value at each number of K.

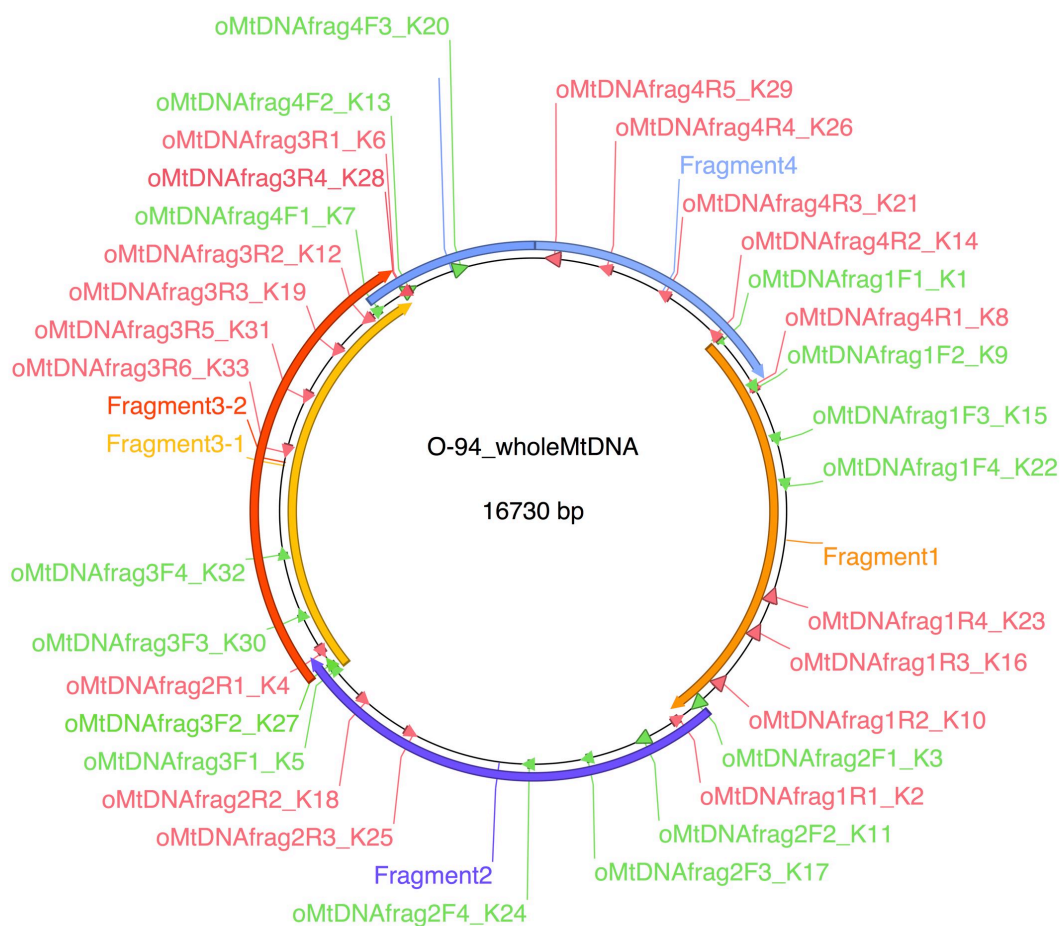


Figure S3: Mitochondrial genome map of "Clade C" showing the primer positions (green arrowheads: forward; pink arrowheads: reverse) and amplified PCR fragments (arrows).

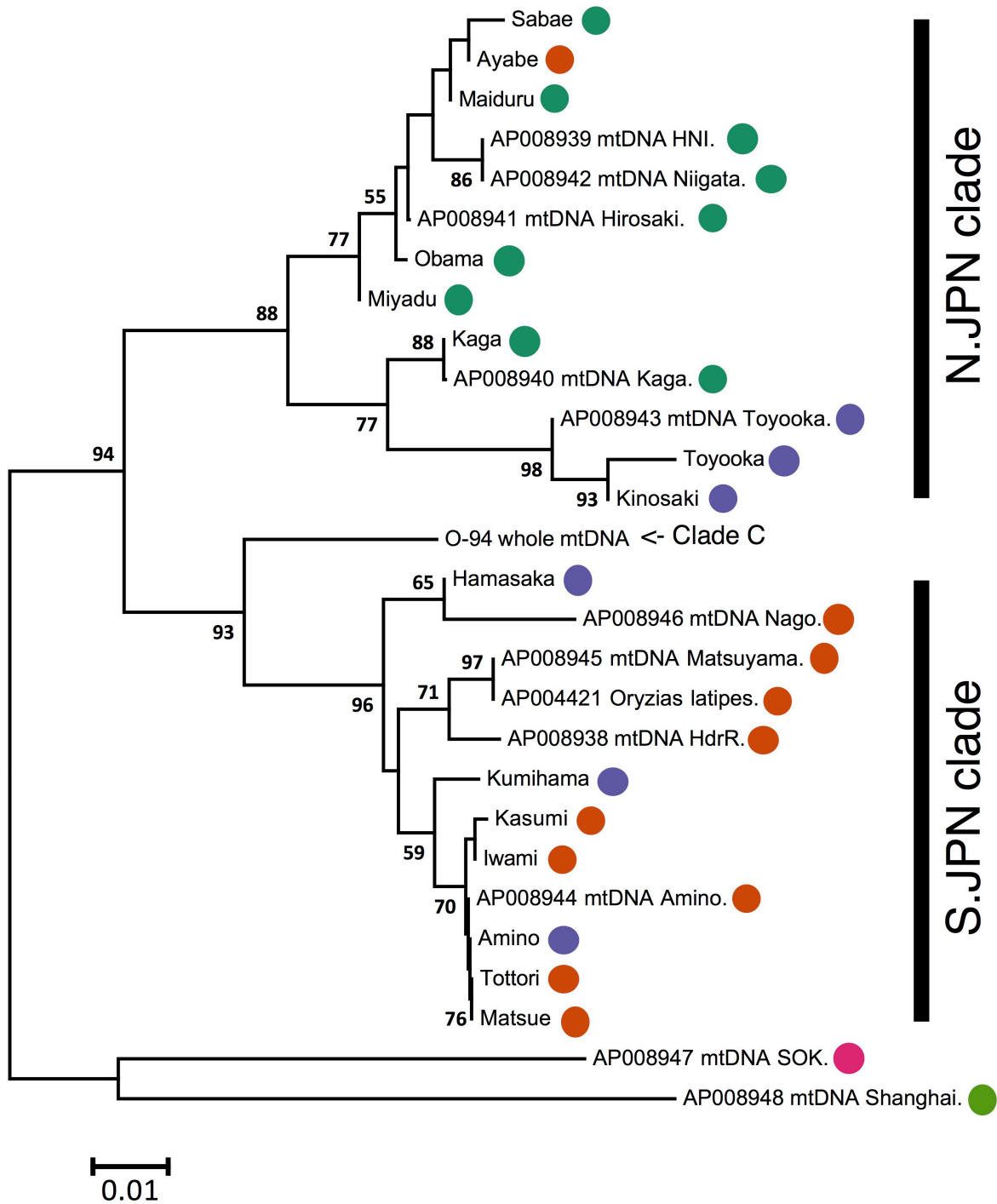


Figure S4: Phylogenetic tree using the neighbor-joining method based on the partial mitochondrial DNA sequences. The numbers starting with "AP" are the accession numbers of each OTU. Each color circle represents the geographical group corresponding to it in fig. 1.

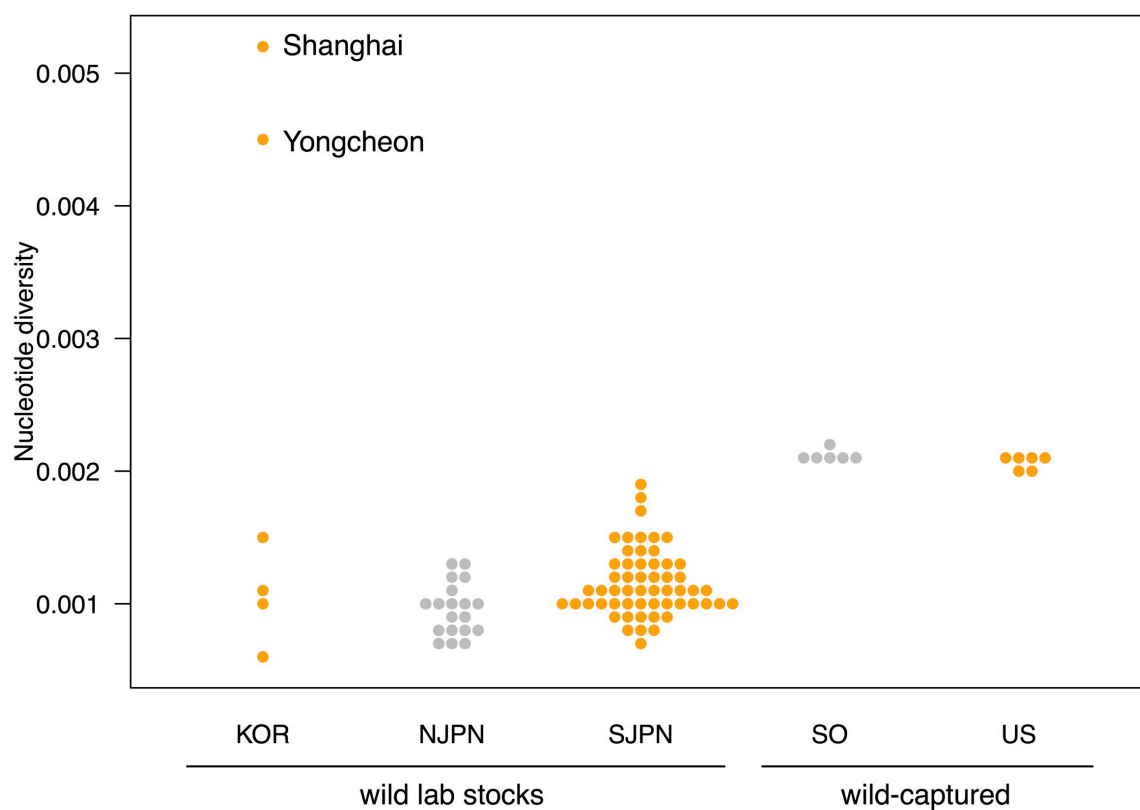
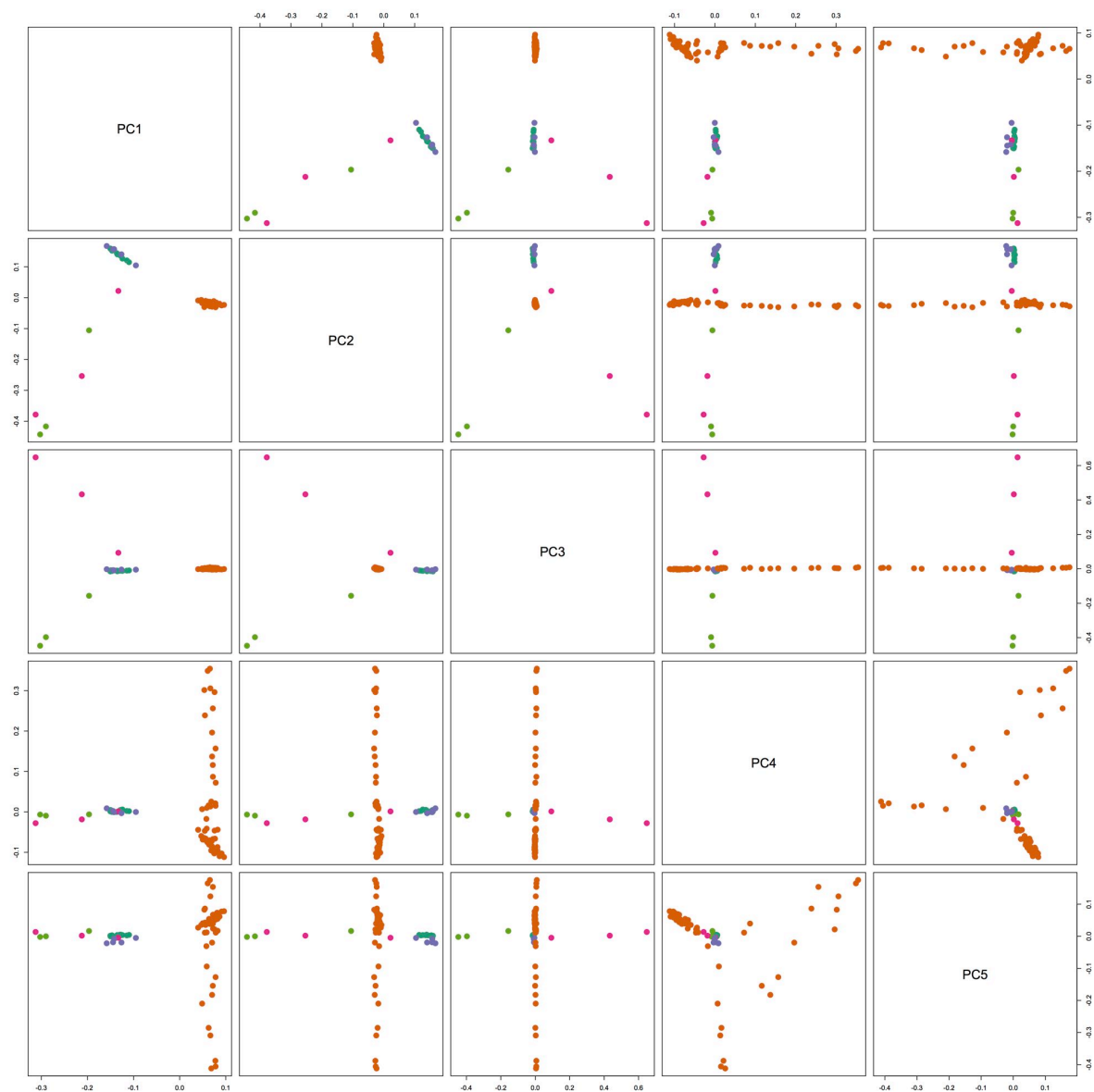


Figure S5: Nucleotide diversity in each individual in Groups.

68
69
70
71



72
73
74
75

Figure S6: Pairwise comparisons of the five principle components (from PC1 to PC5).
Each color circle represents the geographical group corresponding to it in fig. 1 and fig. 2.

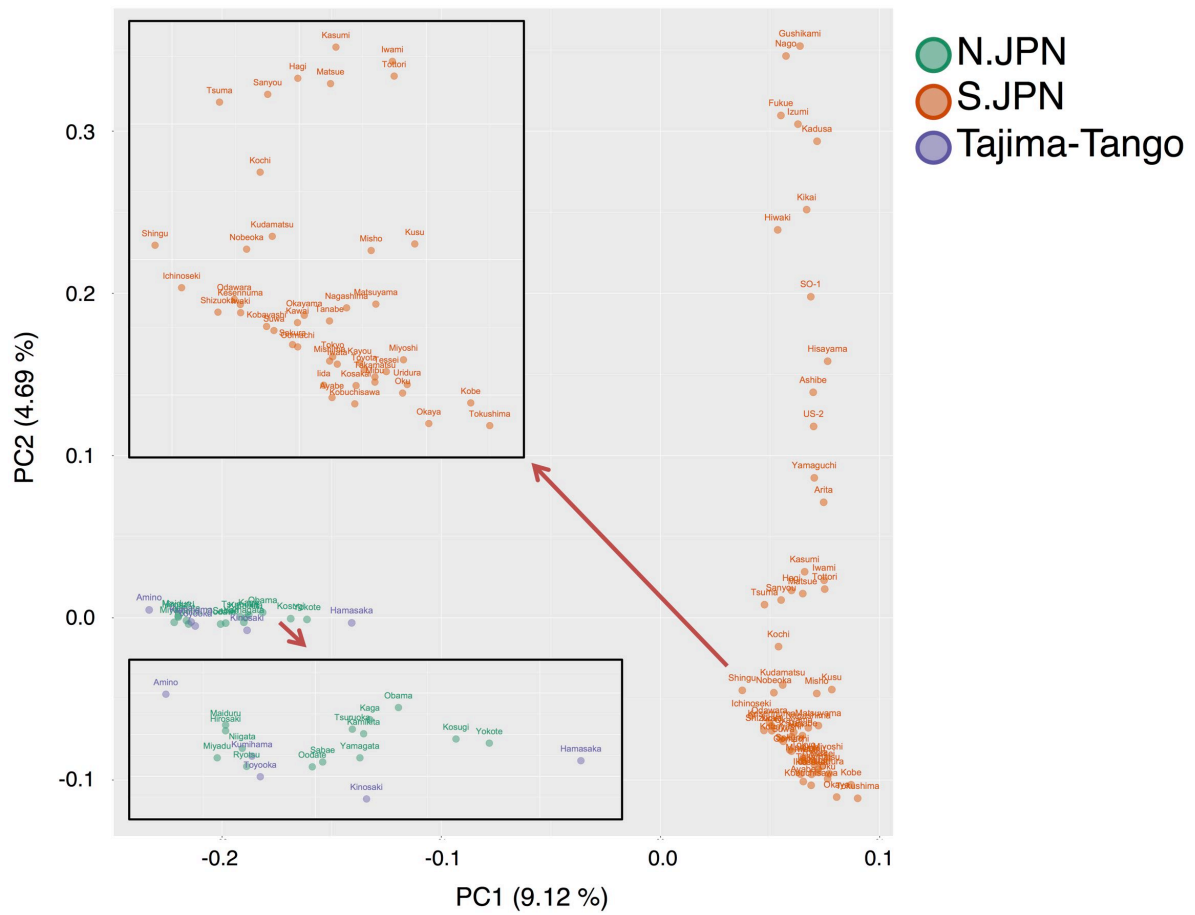
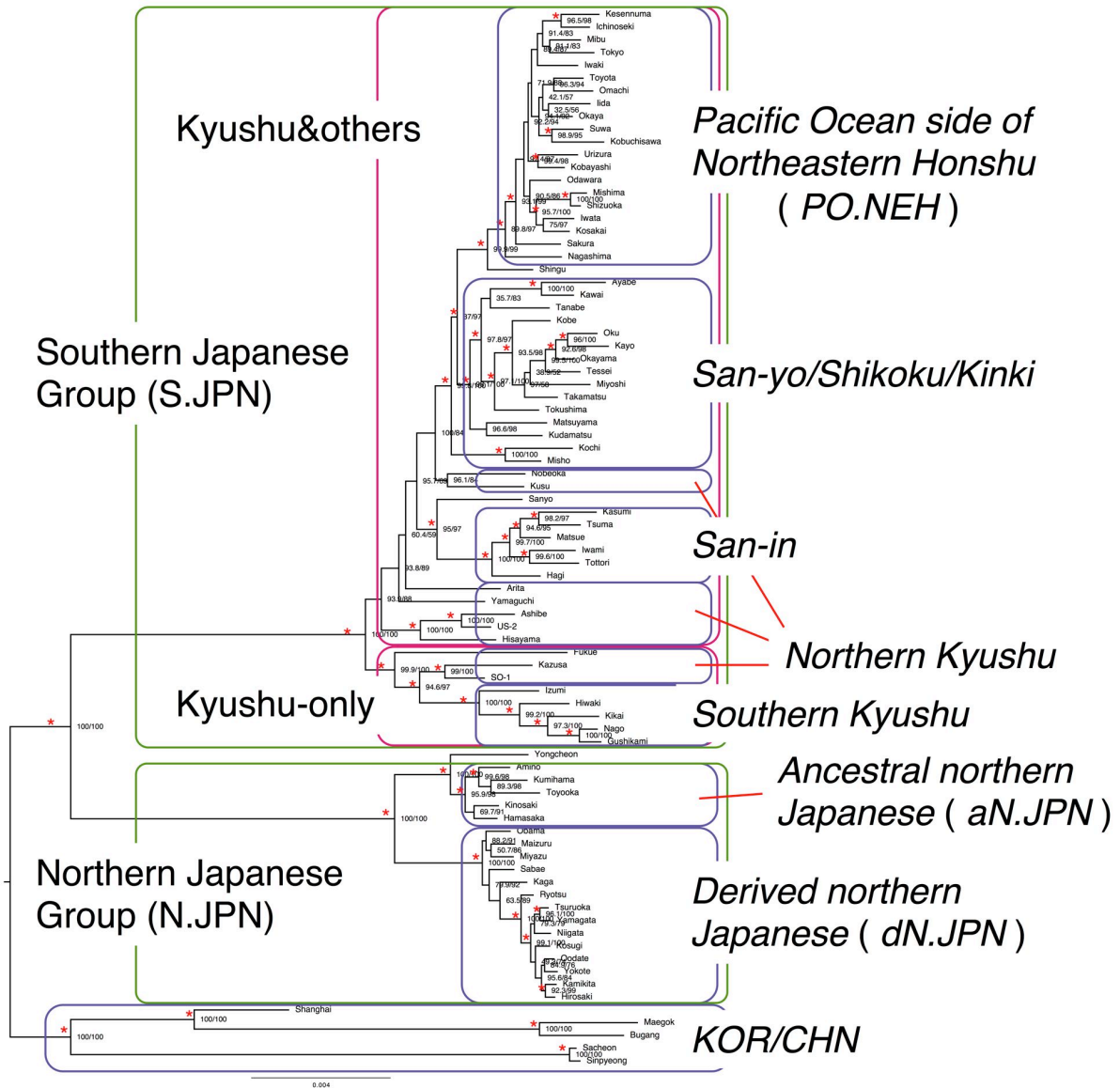


Figure S7: Principle component analysis (PCA) plot using SNPs in the Japanese archipelago with the population names at each point.

84
85



86
87
88
89
90
91
92
93
94
95

Figure S8: Phylogenetic tree using the maximum likelihood method with SH-aLRT (left) and ultrafast bootstrap (right) values as node labels. Typically, SH-aLRT $\geq 80\%$ and UFboot $\geq 95\%$ can be relied upon and are indicated by red asterisks. Re-defined genetic groups by reliable bootstrap values on the ML tree are shown by encircling colored lines and *Italic* fonts. Our lab-stocks are composed of the eight distinct subgroups: PO.NEH, San-yo/Shikoku/Kinki, San-in, Northern Kyushu, Southern Kyushu, Ancestral northern Japanese (aN.JPN), Derived northern Japanese (dN.JPN) and KOR/CHN, based on genome-wide SNPs (see also Discussion).

96
97
98
99

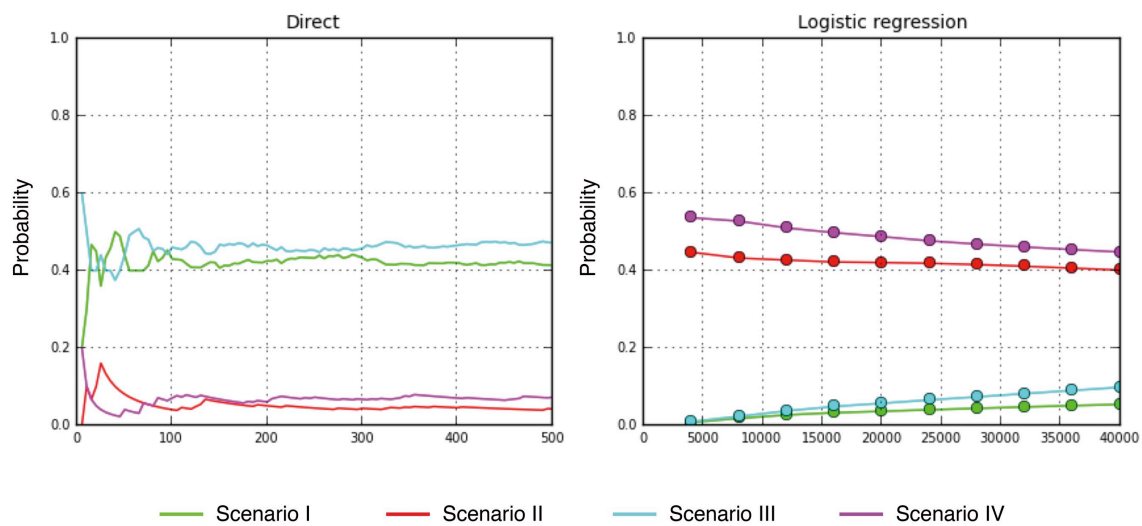


Figure S9: Confidence in scenario choice. These graphs show the four scenario posterior probabilities using the direct comparison approach and logistic regression.

Table S1: Primers for PCR and sequencing to determine the complete mtDNA sequence of Clade C.

Purposes of use	Primer names	Fragment	Length	Sequences (5' ->3')
PCR & Sequencing	oMtDNAfrag1F1_K1	1	27	GCGGGGATAAAAACATAAGACGAGAAG
PCR & Sequencing	oMtDNAfrag1R1_K2	1	28	TAGGCCTAAGAAGTGTGAGGGAAGAAG
PCR & Sequencing	oMtDNAfrag2F1_K3	2	28	GGGCCCTTGGTTTTATTTTCCTCTTTAC
PCR & Sequencing	oMtDNAfrag2R1_K4	2	24	ACATGGGCTTTAGGGAGTCAGAGG
PCR & Sequencing	oMtDNAfrag3F1_K5	3	24	GAAACCAGACAGAGCGGCTTAACG
PCR & Sequencing	oMtDNAfrag3R1_K6	3	25	AGAAAAGGAATTGGGTGAAAGGTCG
PCR & Sequencing	oMtDNAfrag4F1_K7	4	23	CCAACCGGCCTCAATTCAGACTC
PCR & Sequencing	oMtDNAfrag4R1_K8	4	25	CTTAGGTCTTTTGCAATATCCGGGC
Sequencing	oMtDNAfrag1F2_K9	1	26	GGTTCAATTCCTCTCCTTAAGTATGC
Sequencing	oMtDNAfrag1R2_K10	1	27	GAATAGGTGTTGGTAGAGAATAGGGTC
Sequencing	oMtDNAfrag2F2_K11	2	24	TTATGGCACATCCGTCTCAACTAG
Sequencing	oMtDNAfrag3R2_K12	3	23	GTTGCCAATATAAGGGACAGCAG
Sequencing	oMtDNAfrag4F2_K13	4	21	CCTGAATTGGTGGAATGCCTG
Sequencing	oMtDNAfrag4R2_K14	4	23	CAGGCCTTCATTTAAAGGCAAG
Sequencing	oMtDNAfrag1F3_K15	1	20	AACCCTCGCTGAAACCAACC
Sequencing	oMtDNAfrag1R3_K16	1	24	GGGTGCCGATATCTTTGTGATTAG
Sequencing	oMtDNAfrag2F3_K17	2	24	CCTATTGTTGTTGAAGCTGTCCCC
Sequencing	oMtDNAfrag2R2_K18	2	19	TCAGTTCCGTGGGTGCGAG
Sequencing	oMtDNAfrag3R3_K19	3	24	CGAAGGTTGGCCATTAAGTTCTTG
Sequencing	oMtDNAfrag4F3_K20	4	23	CCAAAAAGTCCCAAAAAAGTCCC
Sequencing	oMtDNAfrag4R3_K21	4	24	GGTCTGTCACCCCTACTTGAAGAT
Sequencing	oMtDNAfrag1F4_K22	1	27	GAACCCTACCTGAAGAGATCAAACTC
Sequencing	oMtDNAfrag1R4_K23	1	27	CTTGGTCTCAATACTATCCTAAGCCTC
Sequencing	oMtDNAfrag2F4_K24	2	22	GTCCCACTTTGACTAGCAACCG
Sequencing	oMtDNAfrag2R3_K25	2	24	GGGTAAAAAGAATGAAGCAGTTGC
Sequencing	oMtDNAfrag4R4_K26	4	25	AACTTCCCCTCTCATATGCTACACC
PCR & Sequencing	oMtDNAfrag3F2_K27	3	28	CACTCCTCGTACTTCAGAACTCAACAGG
PCR & Sequencing	oMtDNAfrag3R4_K28	3	29	GAAGTGAAGTAAAGGTCGAAATGTAAGGC
Sequencing	oMtDNAfrag4R5_K29	4	22	GGCTAAGCAAGGTGTTATGGGC
Sequencing	oMtDNAfrag3F3_K30	3	22	ACCCTTGCCAACCTAGCACTCC
Sequencing	oMtDNAfrag3R5_K31	3	20	TCGAGGGACATTACGGGCAG
Sequencing	oMtDNAfrag3F4_K32	3	22	CCCCTGTTCAATTTCTCAACG
Sequencing	oMtDNAfrag3R6_K33	3	22	AGGGTTAAGACCAGGGCTCAGG