4-4-2017

Genetic structure of First Nation communities in the Pacific Northwest

Cris E. Hughes
Department of Anthropology and Carl R. Woese Institute for Genomic Biology, Univeristy of Illinois Urbana Champaign,
hughesc@illinois.edu

Mary P. Rogers
Department of Anthropology, University of Illinois at Urbana-Champaign

Amanda C. Owings
Program in Ecology, Evolution and Conservation Biology, University of Illinois at Urbana-Champaign

Barbara Petzelt
Metlakatla Treaty Office, Metlakatla, British Columbia, Canada

Joycelynn Mitchell
Metlakatla Treaty Office, Metlakatla, British Columbia, Canada

See next page for additional authors

Authors
Cris E. Hughes, Mary P. Rogers, Amanda C. Owings, Barbara Petzelt, Joycelynn Mitchell, Harold Harry, Theresa Williams, Dena Goldberg, Damian Labuda, David Glenn Smith, Jerome S. Cybulski, and Ripan S. Malhi

Genetic structure of First Nation communities in the Pacific Northwest

Cris E. Hughes[1,2], Mary P. Rogers[1], Amanda C. Owings[3], Barbara Petzelt[4], Joycelynn Mitchell[4], Harold Harry[5], Theresa Williams[6], Dena Goldberg[7], Damian Labuda[8], David Glenn Smith[9], Jerome S. Cybulski[10], Ripan S. Malhi[1,2,3]

[1]Department of Anthropology, University of Illinois Urbana Champaign, Urbana, IL 61801, [2]Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana Champaign, Urbana, IL 61801, [3]Program in Ecology, Evolution and Conservation Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, [4]Metlakatla Treaty Office, Metlakatla, British Columbia, Canada, [5]Stswecem'c Xgat'tem First Nation, British Columbia, Canada, [6]Splatsin First Nation, Enderby, British Columbia, [7]University of California at San Francisco Helen Diller Family Comprehensive Cancer Center, [8]Departement de Pédiatrie, Centre de Recherche du Centre Hospitalier Universitaire Sainte-Justine, Université de Montréal, Montréal, QC, Canada, [9]Department of Anthropology, University of California, Davis, [10]Canadian Museum of History, Gatineau, QC, Canada.

Corresponding author: Cris Hughes Department of Anthropology University of Illinois at Urbana-Champaign 109 Davenport Hall 607 S. Matthews Ave Urbana, IL 61801 Phone: 217-333-9694 hughesc@illinois.edu

Issue: 88.4

Running header: Genetic structure in the Pacific Northwest

Keywords: mitochondrial DNA, Y chromosome, D9S1120, first nations, migration

ABSTRACT

This study presents genetic data for nine Native American populations from northern North America. Analyses of genetic variation focus on the Pacific Northwest (PNW). Using mitochondrial, Y chromosomal and autosomal DNA variants, we aim to more closely address the relationships of geography and language with present genetic diversity among the regional PNW Native American populations. Patterns of genetic diversity exhibited by the three genetic systems were consistent with our hypotheses, in that we expected genetic variation to be more strongly explained by geographic proximity than linguistic structure. Our findings were corroborated through a variety on analytic approaches, with the unrooted trees for the three genetic systems consistently separating inland from coastal PNW populations. Furthermore, the AMOVA tests support the trends exhibited by the unrooted trees, with geographic partitioning of PNW populations (F_{CT} = 19.43%, p = 0.010 ± 0.009) accounting for over twice as much of the observed genetic variation compared with linguistic partitioning of the same populations (F_{CT} = 9.15%, p = 0.193 ± 0.013). These findings demonstrate a consensus with previous PNW population studies examining the relationships of genome-wide variation, mitochondrial haplogroup frequencies, and skeletal morphology with geography and language.

INTRODUCTION

Demographic events and processes impact the genetic structure of populations; these can include migrations, differential mating opportunities often alongside cultural and language transmission and patterns of exogamy (patrilocal versus matrilocal), and subsistence transitions (Hunemeiere et al., 2012; Kemp et al., 2010; Bolnick et al., 2006; Eshleman et al., 2004). The extent to which these phenomena have influenced genetic variation among indigenous peoples of

the Americas today can be inferred through comparative analysis of population genetic structure and diversity. A variety of genetic systems including autosomal, mitochondrial, and Y chromosomal data have previously been studied (e.g., Wang et al., 2007; Dulik et al., 2012; Tamm et al., 2007; Reich et al., 2012), each contributing to the understanding of population history across the Americas.

Such genetic information is also useful for comparing to population history inferences produced from other non-biological systems such as linguistic and cultural data. In the Pacific Northwest (PNW), historical linguistic studies suggest extensive interactions between the coastal and the interior communities of British Columbia, based on the geographic distribution and inferred linguistic relationships. For example, the Salishan language family is believed to have originated on the Northwest Coast and spread to the interior across the Coastal Range (Elmendorf, 1965; Suttles and Elmendorf, 1963). Linguistic evidence of a Salishan homeland is based on the greater linguistic diversity among the Salishan languages located near the coast than among the Salishan languages on the interior. Kinkade (1991) provided lexical evidence of a coastal origin for the Salishan language family by reconstructing proto-Salish words for several Pacific coastal shellfish and plants. In addition, Verdu and colleagues (2014) completed a genome-wide analysis of PNW populations, and identified substantial genetic differentiation between coastal and interior groups. They suggest a model for a shared origin for the PNW populations, and, after the initial peopling of the region, divergence due to isolation and drift. A scenario of common origin, followed with isolation and drift also fits paleoanthropometric data for this region as morphological analyses of ancient skeletal remains from the coast and interior differ (Cybulski, 2010). Additional investigations in or near the PNW region have heavily focused on mitochondrial DNA (mtDNA) data. Malhi and colleagues (2004) identified patterns

of haplogroup variation between inland and coastal groups, characterized by high frequencies of haplogroup A along the coast, with a marked decrease in frequency moving inland. This evidence, along with shared hypervariable region I (HVS1) sequences among Northwest and California coastal populations (Eshleman et al., 2004), suggests that coastal gene flow together with limited gene flow between coastal and inland populations occurred.

In this study, we aim to more closely address the impact of geography and language on present genetic diversity in the PNW. For the present study, PNW refers to the geography encompassing both the Northwest Coast and Plateau (Cybulski, 2006). We report genetic data from First Nation communities, located in both coastal and inland regions, and with linguistic ties to Salishan and Tsimshianic language families. This regional data, along with data from three additional study samples (Cree, Chipewyan, and Dogrib) and previously analyzed Native American samples from Northern North America, represent adequate linguistic and geographic sampling for observing geographic and linguistic patterns of genetic variation in the PNW (e.g. coastal *versus* inland; southern to northern PNW).

In particular, populations that speak Salish languages can be tested for the relationship of genetic diversity with language and geography, as Salishan family languages are spoken in both coastal (Bella Coola) and inland (Splatsin, and Stswecem'c) geographic locations. Drawing from previous work (Eshleman et al., 2004; Malhi et al., 2004), we hypothesize that geographic proximity will more accurately reflect the genetic variation of the PNW as compared with linguistic affiliations. In order to compare our results with previous findings (Malhi et al., 2004; Eshleman et al., 2004) we employed mtDNA HVS1 data for this test.

Additionally, we generated Y-chromosomal data to infer whether patterns of mtDNA genetic variation are mirrored by paternal lineage variation. Colonialism has been shown to

differentially impact the uniparental DNA systems, with Y chromosomes exhibiting a much higher degree of non-indigenous influence relative to the mtDNA system (e.g., Bolnick et al., 2006). Because the present study focuses on linguistic and geographic effects on genetic diversity in a pre-European contact context, uniparental markers are useful in that those individuals with haplogroups exclusive to the Americas can be identified and used for analyses that specifically address Native American genetic diversity. We hypothesize 1) that the Y-chromosomal data will yield fewer autochthonous haplogroups than the mtDNA data, and 2) that Y-chromosomal diversity will mirror patterns of mtDNA diversity, as geographic proximity has been shown to more accurately reflect genome-wide variation, as in Verdu et al. (2014).

Lastly, we explore patterns of variation for a private (specific to the Americas) autosomal genetic variant, the 9-repeate allele of D9S1120 (Schroeder et al., 2007; Schroeder et al., 2009). Previous work on genome-wide variation in the PNW (Verdu et al., 2014) identified genetic differentiations between coastal and inland groups. Therefore, we hypothesize that allelic variation of D9S1120 will conform to the geographic pattern similarly hypothesized for the uniparental markers. The combination of mtDNA HVS1, Y-chromosomal DNA, and an autosomal genetic variant provides a multifaceted approach to investigating the relationship of geography and language with population genetics of the PNW region.

This study contributes to an understanding of genetic variation at the regional-level within the context of local population histories, specifically the geographic and/or linguistic patterning of genetic diversity in the PNW. Furthermore, the findings presented here expand the understanding of present-day genetic variation in the PNW region, which is an area that has previously lacked adequate coverage in continental-wide studies of Native American genetic diversity.

MATERIALS AND METHODS

Study Populations

This study included generation of novel data from individuals from First Nations communities in the Pacific Northwest with immediate ancestors from: Coast Tsimshian (Metlakatla and Lax Kw'alaams), Laxgalts'ap, Splatsin, and Stswecem'c. Data from three additional groups from northern North America were generated (Cree, Chipewyan, and Dogrib). Participating tribal members consented to participate in the study and sampling protocols were approved by Institutional Review Boards at the University of Illinois, Urbana-Champaign (submitted by author RSM) and University of Montreal (submitted by author DL).

Genetic data was generated to adequately address each of the research goals outlined in the introduction, which required that linguistic and geographic (coastal and inland) variation be represented in the sample. Table 1 displays the tribal affiliation, linguistic group, and sample size for all study samples, as well as comparative samples drawn from published data for mtDNA and Y-chromosomal analyses (Lorenz and Smith, 1996; Ward et al., 1993; Schurr et al., 2012; Johnson and Lorenz, 2006; Kemp et al., 2010). The map in Figure 1 displays the geographic location for all samples included in this study. The previously published samples included in the present study were selected based on their geographic and linguistic affinities. If linguistic or geographic trends emerge in the mtDNA variation for the PNW regional samples, it is useful to situate these findings within the broader North American geographic genetic trends to assist in interpreting the findings. For example, while linguistically different, the Yakama and Splatsin both occupy territories within the Northwest Plateau. The genetic proximity of these groups will in turn contribute to inferring whether genetic variation among the study's PNW groups is more linguistically or geographically structured.

Methods

The following methods were performed on 105 PNW samples, from Coast Tsimshian, Splatsin, Laxgalts'ap, and Stswecem'c. DNA was collected from saliva using buccal swabs or Oragene© saliva kits. DNA from buccal swabs was extracted using the protocol adapted from Miller et al. (1988). DNA from Oragene© saliva kits was extracted by first incubating the samples at 50° C for two hours. We then added 20 μL of Oragene© DNA Purifying Solution to 500 μL of sample, which was put on ice for 10 minutes and centrifuged for five minutes at 13,000 RPM. The supernatant was purified with two ethanol washes, the first with 500 μL of 100% ethanol and the second with 250 μL of 70% ethanol. Extracted samples were kept at -20°C as a stock solution with sample dilutions created at a concentration of 10 ng/μL.

Genotyping and sequencing mtDNA

The following methods were performed on sample dilutions from Coast Tsimshian, Splatsin, Laxgalts'ap, Stswecem'c, Dogrib, Cree, and Chipewyan (N = 188). The DNA samples were amplified with 15 μL polymerase chain reactions (PCRs) containing: 1.5 μL 10x buffer, 0.6 μL 10mM dNTPs, 0.6 μL 50mM $MgCl_2$, 0.6ul 5uM forward primer, 0.6 μL 5uM reverse primer, 0.12 μL Platinum® Taq Polymerase, 1.5 μL 50% glycerol, 7.48 μL water, and 2 μL 10 ng/μL DNA. Primers used were designed for the HVSI of the mitochondrial control region (Malhi et al., 2007; Kemp et al., 2006). Final DNA concentrations in the PCRs were 1.33 ng/μL. PCRs included initial denaturation at 94°C for 2 minutes, followed by 40 cycles of denaturing (15 seconds at 94°C), annealing (30 seconds at 54°C), extension (1 minute at 72°C), and a final elongation of 5 minutes at 72°C. PCR product sizes were determined by separating 5 μL of

amplicons on 2% agarose gels stained with ethidium bromide and compared to a size standard ladder.

The amplified DNA was then purified with USB® ExoSAP-IT® following the manufacturer's protocol. DNA was sequenced at the UIUC Core Sequencing Center on an ABI 3730xl DNA analyzer. Electropherograms were visually inspected using the program Sequencher® 4.8 (Gene Codes Corp., 2007), and aligned to the revised Cambridge reference sequence (Andrews et al., 1999). Single nucleotide polymorphisms (SNPs) were used to cluster the individuals into haplogroups based on the defining mutations described by Tamm et al. (2007).

Y-STR and Y-chromosomal DNA

The following methods were performed on a subset (*N* =57) of the PNW male samples for Coast Tsimshian, Splatsin, Laxgalts'ap, and Stswecem'c. Y-chromosome haplogroups were determined for all samples. Allelic discrimination for haplogroups Q, C and R was first performed via the Applied Biosystems 7900HT Fast Real-Time PCR system following the TaqMan® protocol. Primers were used with a normalized reporter specific to mutations found in haplogroups C (M130), Q (M242), and R (M173) (Karafet et al., 2008). Haplogroup Q samples were further analyzed with a normalized reporter specific to haplogroup Q-M3 (M3). For haplogroups C, R, Q, and Q3, SNP database (dbSNP) reference SNPs (rs) 2032666, rs8179021, rs3212294 and rs2032624 were used (http://www.ncbi.nlm.nih.gov/SNP/). Detection of a particular SNP after fluorescent primer annealing was determined using SDS 2.3 PCR Software Analysis (Applied Biosystems, 2007). The presence of an annealed primer indicated

the haplogroup to which the sample belongs. Samples that annealed properly were later compared to the Y-STR haplogroup prediction.

Y chromosomal DNA haplogroups generated from TaqMan® analyses were confirmed using the AmpFlSTR© Yfiler© PCR Amplification Kit. Seventeen Y-STR loci were analyzed and read using Genemapper version 4.0. Haplogroups were predicted using the Whit Athey online predictor tool using the following settings: FTDNA order and 27 haplogroups (http://www.hprg.com/hapest5/hapest5b/hapest5.htm).

D9S1120

The following methods were performed on a subset of the PNW samples (*N*=119), from Coast Tsimshian, Splatsin, Laxgalts'ap, and Stswecem'c. Primers for the D9S1120 locus were drawn from Schroeder et al. (2007). The forward primer was labeled with NED™ dye (Applied Biosystems) for STR genotyping. The DNA samples were amplified with 29 μL PCRs containing: 5 μL of 10ng/μL DNA, 2.5 μL 10x buffer, 1 μL 10mM dNTPs, 1 μL 50mM $MgCl_2$, 13.3 μL water, 5 μL 5% glycerol, 1 μL 20uM primers, and 0.2 μL Platinum® Taq Polymerase, with final concentrations of 1.78 ng/μL DNA in PCRs. Thermocycler settings included an initial denature step at 95°C for 11 minutes, followed by 40 cycles of denaturing (45 seconds at 95°C), annealing (1 minute at 59°C), extension (70 seconds at 72°C), and a final elongation of 60 minutes at 65°C. Length variations of amplicons were determined using an ABI 3730xl Genetic Analyzer. The software Genemapper version 3.7 was used for fragment analysis with a size standard of GeneScan™LIZ500™ (Applied Biosystems, 2004). The 9RA allele product size was determined to be 168 base pairs in length (Phillips et al., 2008). Each tetra-nucleotide repeat

thereafter was considered the consecutive repeat allele in concurrence with previously published data (Philips et al., 2008).

Data analysis

MtDNA sequences were aligned in BioEdit and haplogroup frequencies were calculated in Microsoft Excel. Individuals with non-autochthonous haplogroups were excluded from the following analyses. For the 18 population samples, Arlequin was used to test pairwise differences in haplotype frequencies using a modified Fisher's exact test of population differentiation for an expanded number of populations, and the fixation index F_{ST} was generated for all population pairs. A principal coordinates plot of the F_{ST} matrix was generated in GenAlEx 6.5 (Peakall and Smouse, 2012). To focus more closely on the genetic diversity among PNW regional samples, an additional pairwise F_{ST} matrix was generated in Arlequin for 11 of the 18 sampled populations, and served as the basis for creating the Fitch-Margoliash Least Squares derived unrooted tree using the program PHYLIP v.3.6 (Felsenstein, 2005). For statistical comparison (described below) with the Y-chromosomal F_{ST} matrix, a third pairwise F_{ST} matrix was generated using mtDNA data for the same six PNW regional samples (Kaigani Haida, Coast Tsimshian, Laxgalts'ap, Splatsin, and Stswecem'c, and Tlingit) utilized in the Y-chromosomal analysis.

To assess the amount of genetic variation captured in geographic and linguistic groupings of the PNW populations, two analyses of molecular variance (AMOVA) were completed for mtDNA haplotype variation, one which grouped regional samples into inland (Splatsin and Stswecem'c) *versus* coastal (Bella Coola, Coast Tsimshian and Laxgalts'ap) groups, while the

second AMOVA was run with Tsimshianic language speakers (Coast Tsimshian and Laxgalts'ap) *versus* Salish language speakers (Bella Coola, Splatsin and Stswecem'c).

Y-chromosomal haplogroup frequencies were calculated for PNW populations, with individuals belonging to non-autochthonous haplogroups (i.e., I1, I2b1, J1, L, G2a) excluded from the following analyses. A pairwise F_{ST} matrix was generated in Arlequin and served as the basis for generating the Fitch-Margoliash Least Squares derived unrooted tree using the program PHYLIP v.3.6 (Felsenstein, 2005). A Mantel test was performed to test the hypothesis of correspondence in population differentiations among the six PNW populations using the mtDNA and Y-chromosomal F_{ST} matrices.

Allele frequencies of the D9S1120 for the PNW samples were calculated and compared with published allele frequencies for North American tribes (Schroeder et al., 2007). Individuals with non-autochthonous Y-chromosomal and/or mtDNA haplogroups were not included in the following analyses. A pairwise F_{ST} matrix was generated using Arlequin and served as the basis for generating the Fitch-Margoliash Least Squares derived unrooted tree using the program PHYLIP v.3.6 (Felsenstein, 2005).

RESULTS

Mitochondrial DNA Variation Trends: Language and Geography

The mtDNA haplogroup frequencies for the PNW and northern North American regional population samples are presented in Table 2. Only the Splatsin exhibit a non-autochthonous haplogroup (i.e., haplogroup H at 5%). Among the five Native American haplogroups exhibited by the study samples, haplogroup A is generally in greatest frequency, with the exception of the Splatsin and Stswecem'c, both inland groups.

The population differentiation tests were significant ($\alpha = 0.05$) for all sampled populations, thus rejecting the null hypothesis of panmixia (for all pairwise samples, $p = 0.000 \pm 0.000$). The principal coordinates plot (Figure 2) of the mtDNA pairwise F_{ST} matrix for regional and continental comparative samples (Table 3) accounted for 90.39% of the variation in the first and second dimensions. Three general clusters for the PNW populations are notable: interior, transitional (i.e. from interior to coastal), and coastal clusters, and correspond to general geographic locations of the PNW populations, with the plotted data progressing from southern interior to northern coastal and finally to northern interior. The interior cluster includes the Splatsin, which is the most southern and interior community of the PNW sampled here, and occupies the lower left of the plot, associated with other interior tribes close to the U.S.-Canadian border (Yakama and Northern Paiute). The transitional cluster includes other interior tribes that are closer to the coast and north of the Splatsin (e.g. Stswecem'c), as well as coastal tribes that are located along the southern coast of the PNW (e.g. Nuu-chah-nulth and Bella Coola). The southern coastal locations of the Nuu-chah-nulth and Bella Coola means that these two populations are in closer proximity to the inland PNW populations sampled in this study, a feature mirrored in the Nuu-chah-nulth and Bella Coola's proximity on the plot to the inland Stswecem'c.

The coastal cluster includes PNW tribes located on the coast, with those to the north positioned on the right portion of this cluster (Laxgalts'ap, Kaigani Haida, Haida Gwaii). Additionally, the Tlingit, the most northerly located coastal population included in the study is positioned even further to the right in the plot, representing the transition point between the northern coastal and interior Canadian tribes (Dogrib and Chipewyan).

The Fitch-Margoliash Least Squares derived unrooted tree (Figure 3) conforms to a similar pattern of geographic-based associations. Observing the geographic patterning of the tree from left to right, the samples generally follow a southern interior (Yakama, Stswecem'c, and Splatsin) branching, followed by the main trunk progressing from southern coastal (Nuu-chah-nulth), to coastal (Bella Coola, Laxgalts'ap, Coast Tsimshian, Kaigani Haida, Haida Gwaii), to northern coastal (Tlingit) and ending with the northern interior (Dogrib).

Mitochondrial DNA results do not exhibit as strong a pattern for linguistic structure for the sampled populations with Salishan language family associations (Stswecem'c, Splatsin and Bella Coola). In the principal coordinates plot (Figure 2), Bella Coola is more closely positioned to other coastal or near-coastal tribes, such as Nuu-chah-nulth and Coast Tsimshian. Furthermore, the two inland Salishan family tribes, Splatsin and Stswecem'c, exhibit the lowest haplogroup A frequencies, while the coastal populations all exhibit haplogroup A in greater frequencies than any other haplogroup. The AMOVA tests corroborate the trends exhibited by the principal coordinates plot. The AMOVA test with geographic partitioning of PNW populations (F_{CT} = 19.43%, p = 0.010 ± 0.009) accounts for more than twice as much of the observed haplotype variation compared with linguistic partitioning of the same populations (F_{CT} = 9.15%, p = 0.193 ± 0.013).

Geographic Patterning of Y Chromosomal and Autosomal DNA

All Y-chromosomal DNA haplogroups inferred from the predictor tool matched with the data generated from TaqMan® analyses and haplogroup frequencies are presented in Table 4. The two inland sampled populations, Stswecem'c and Splatsin, exhibit greater frequencies of non-autochthonous haplogroups as compared to the two coastal groups (Coast Tsimshian and

Laxgalts'ap), which suggests that European admixture in the PNW is geographically structured. In general, Y-chromosomal results exhibit greater frequencies of non-autochthonous haplogroups as compared to mtDNA haplogroups. Specifically, while only one of the four PNW sampled tribes exhibited non-autochthonous mtDNA haplogroups, three of these four tribes exhibited non-autochthonous Y-chromosomal haplogroups, with their cumulative frequencies (Haplogroups R1b, I1, I2b1, J1, L and G2a) ranging from 0.44 to as high as 0.58 in the four PNW sampled populations.

To relate Y-chromosomal Native American haplotype diversity to the geographic patterning of mtDNA native haplogroups, the relationships of the PNW tribes on the Fitch-Margoliash Least Squares derived unrooted trees is useful to observe (Figure 4). For the unrooted Y-chromosomal tree, the inland Stswecem'c and Splatsin are branched together, and separated from the coastal tribes, Laxgalts'ap and Coast Tsimshian. This distribution of inland versus coastal tribes is in agreement with the mtDNA data as seen in the principal coordinates plot and similarly generated tree (Figure 2 and 3). The Mantel test for correlation among the mtDNA and Y-chromosomal F_{ST} matrices was non-significant ($R^2 = 0.06$, $p = 0.210$). The non-significance may derive from the small sample sizes utilized for this test, or may derive from different rates of historic gene flow among the tested populations for males and females.

D9S1120 allele frequencies for the PNW samples (Stswecem'c, Splatsin, Laxgalts'ap and Coast Tsimshian) are presented in Table 5. The 9RA allele is present in this region, ranging in frequency from 3-18%. The Stswecem'c, Laxgalts'ap and Coast Tsimshian 9RA frequencies fall within the range of those previously documented for North and South American populations, from 10-97%, but all four PNW study sample frequencies are lower than the average for the rest of the Americas at 31.7% (Schroeder et al., 2007). At 3%, the Splatsin exhibit a 9RA frequency

approximately five times less than the other three PNW sampled populations as well as three times less than the lowest recorded Native American populations sampled by Schroeder et al. (2007), the Seri at 10%. When focusing on the PNW regional variation among the four tribes, the unrooted tree for D9S1120 allelic variation (Figure 5) conforms to the previously established patterns for mtDNA and Y-chromosomal data, with the inland Stswecem'c and Splatsin branched together and separated from the coastal tribes (Laxgalts'ap and Coast Tsimshian).

DISCUSSION

Mitochondrial DNA Trends in the Pacific Northwest: Geographic and Linguistic Considerations

Geography and language are examined as plausible factors influencing the patterns of mtDNA genetic variation exhibited among the PNW populations. Mitochondrial DNA haplogroup A frequencies indicate a marked difference among coastal and inland populations, a trend which can be visualized in the unrooted tree (Figure 3). The decrease in haplogroup A towards the interior reflects a previous trend identified in coastal and inland populations of California and Oregon (Eshleman et al., 2004; Malhi et al., 2004), and is consistent with the overall model inferred by Verdu et al. (2014), with a shared origin for the Pacific Northwest populations, and, after the initial peopling of the region, divergence due to isolation and drift.

The principal coordinates plot (Figure 2) suggests a slightly less simplistic interpretation of these relationships, as the inland Stswecem'c is equidistant to both the inland Splatsin and one southern coast PNW tribe, the Nuu-chah-nulth. However, this positioning may be less related to the relationship with each other, and more related to the inclusion of whatever mtDNA diversity is shared with the inland groups the Cree and Sisseton/Wahpeton, as they are

also closely positioned to the Nuu-chah-nulth and Stswecem'c on the plot. This is supported by the mtDNA unrooted tree primarily comprised of PNW populations (thus not including the Cree or Sisseton/Wahpeton), which situates the Splatsin and Stswecem'c on their own branch, separate from the PNW coastal tribes, including the Nuu-chah-nulth.

Explanations for the genetic proximity of the Splatsin and Stswecem'c can be a result of both geographic and linguistic processes. The Splatsin and Stswecem'c represent the two most interior and southern of the PNW populations used in the Y-STR and D9S1120 analyses, and are thus presumed to have increased opportunities for contact and subsequent gene flow with Europeans and East Asians (Verdu et al., 2014). Additionally, the Stswecem'c and Splatsin both speak languages derived from the Interior Salishan language family, suggesting a common linguistic (and potentially genetic) ancestor. In contrast, the Bella Coola also speak a Salishan-derived language, but are not as closely affiliated with the Stswecem'c and Splatsin in the mtDNA analyses as they are with coastal PNW populations (Figure 2 and 3). This comparison suggests that for the three Salishan language family populations present in this study, language does not appear to have as significant of an impact on the exhibited genetic variations as compared to geographic proximity.

Consideration of the expanded regional analysis of the mtDNA data further demonstrates the discrepancy between linguistic and geographic patterns of genetic similarity. As with previous comparisons of linguistic and mtDNA genetic diversity among Native North American populations (Hunley and Long, 2005), the genetic variation observed in the PNW populations does not explicitly correspond with linguistic families (see Table 1 for linguistic associations). For example, when observing the principal coordinates plot distributions of populations that speak languages of the two most prevalent language families included in this

study (Salishan and Tsimshianic), their positionings are more clearly explained by geographic than linguistic structure. Specifically, the Laxgalts'ap, who speak Nisga'a, a language included in the Tsimshianic language family, are tightly clustered with other coastal populations, which speak languages from the Na-Dené linguistic family (as proposed by Sapir [1915], including the Tlingit sample) and those speaking Haida isolates (Haida Gwaii and Kaigani Haida samples).

Additionally, the three Salishan speaking populations (Splatsin, Stswecem'c and Bella Coola) are dispersed along the first principal coordinate, with the Bella Coola being the most proximate of the three to the coastal cluster, while the Splatsin population is situated near other southern interior populations including the Yakama and Northern Paiute. The proximity of the Yakama and Splatsin is notable because while the Yakama speak a language affiliated with the Plateau Penutian family, both groups occupy territories within the Northwest Plateau, further confirming the geographic model of genetic similarity among the regional groups.

Lastly, the plotting of the Tlingit is also suggestive of a geographic association with the mtDNA genetic variation, in that it appears to represent an intermediate position between coastal and northern inland groups, which mimics this groups geographic positioning (Figure 1). While the Tlingit are the northernmost sample affiliated with the coastal cluster of the plot, oral histories of the Coast Tsimshian and Tlingit document inter-village interactions among these groups, which often entailed slave-raiding where women and children were taken captive (Mitchell, 1984), which potentially increased gene flow among the Tlingit and its southern coastal counterparts. This history of gene flow with the more southerly located PNW tribes is corroborated with the principal coordinates plot (Figure 2), with the Tlingit in close proximity to the cluster of southern PNW coastal populations (Laxgalts'ap, Kaigani Haida and Haida Gwaii).

The AMOVA analyses of mtDNA haplotype variation agree with the geographically patterned diversity exhibited in the unrooted tree and principal coordinates plot. Specifically, the geographic groupings explain over twice as much genetic variation as compared to the AMOVA results performed with linguistic groupings (19.43% *versus* 9.15%). The samples available for comparison in the AMOVA analyses were limited to speakers of Salishan or Tsimshianic languages, therefore the interpretation here as geography explaining more genetic diversity than linguistic affiliation cannot be applied to populations beyond those present in the AMOVA analyses.

Comparison across genetic systems

The greater frequency of non-autochthonous haplogroups in Y-chromosomal results compared to the mtDNA results potentially reflects post-contact gene flow from European males into Native American populations, an occurrence which has been documented in other regions North America as well (Karafet et al., 1999; Bolnick et al., 2006). In the present study, mtDNA haplogroups for all analyzed samples are almost exclusively (97%) comprised of founding haplogroups (A, B, C, D and X). Only the Splatsin exhibit mtDNA haplogroups with possible origins tracing to Europe (Achilli et al., 2004), and only represents 5% of the haplogroup variation present in that sample.

For sampled individuals with Native American mtDNA and/or Y chromosomal haplogroups, there is general agreement of the genetic relationships for the two uniparental systems, as depicted in similarities of the unrooted trees (Figures 3 and 4) where the coastal and inland samples are separated. Furthermore, the corroboration of the allelic variation of D9S1120 (Figure 5) with the geographic patterning of the uniparental results suggests that European

admixture may not significantly affect the trends among the three genetic systems, as individuals with only non-Native American uniparental haplogroups were excluded from the D9S1120 analyses, and thus potentially still included individuals with autosomal non-native genetic variation.

Continental trends for genetic variants of D9S1120

The 9RA is present in all PNW sampled populations, corroborating previous findings (Schroeder et al., 2007; Zhivotovsky et al., 2003) on the ubiquity of this allele within the Americas, which had not been previously documented in this region. All but the Splatsin sample frequencies (3%) are within the lower end of the range (10%) previously estimated across the Americas (Schroeder et al., 2007). Because the 9RA is considered a private allele in the Americas, the low frequencies in Splatsin may be a result of European admixture, as evidenced by the comparably high percentage of non-autochthonous Y-chromosomal haplotypes (R1b at 55%). However, Verdu and colleagues' (2014) analysis of over 600,000 genome-wide SNPs for the same Splatsin individuals analyzed in the present study indicated that Splatsin have some of the lowest amounts of non-indigenous ancestry for the PNW populations tested in their study. These findings suggest that an alternative to historic intercontinental gene flow may explain the low 9RA frequencies in Splatsin, possibly genetic drift. Furthermore, when those Splatsin individuals with non-autochthonous Y-chromosomal and/or mtDNA haplogroups were excluded from the D9S1120 allele frequency analysis, the increase in 9RA frequency was only marginal, from 3% to 5%. Regardless of the 9RA frequencies, when all alleles for D9S1120 were used to generate the unrooted tree for the PNW samples (Figure 5), we see geographic structure similar

to the Y-chromosomal and mtDNA trends, with the coastal populations (Coast Tsimshian and Laxgalts'ap) clustering apart from the inland populations (Splatsin and Stswecem'c).

Conclusions

We hypothesized that genetic variation of the PNW study populations would be better explained by geography than linguistic structure. Our findings demonstrate a consensus across a variety of genetic systems for sampled populations on the Pacific Northwest coast, with the coastal and inland populations being more genetically similar within these regions than to each other. These patterns of geographic structure are consistent with previous studies examining genetic variation (Verdu et al., 2014; Malhi et al., 2004; Eshleman et al., 2004) and skeletal variation (Cybulski, 2010) among PNW populations. One area in which we hypothesized divergent results for the mtDNA and Y-chromosomal data was in the frequency of non-autochthonous haplogroups. Consistent with previous investigations on additional North American populations (Bolnick, 2006; Karafet et al., 1999), the PNW populations in the present study exhibited more non-autochthonous haplogroups for Y-chromosomal than in mtDNA.

The mitochondrial, Y-chromosomal and autosomal DNA findings of the present study provide a clearer understanding of genetic diversity at the regional level within the context of local population histories. Building on previous mtDNA (Eshelman et al., 2004; Malhi et al., 2004) data that identified the coastal-inland diversity along the west coast of North America, the present study introduced additional PNW populations and new genetic data (Y-chromosome STR and autosomal STR data) to expand the regional comparison. These findings not only inform on local PNW histories, but fill in the gaps of continental-wide studies of Native American genetic diversity that lacked adequate coverage for northwest North America.

Received 6 January 2016; accepted 11 October 2016.

LITERATURE CITED

Achilli, A., C. Rengo, C. Magri et al. 2004. The Molecular Dissection of mtDNA Haplogroup H Confirms that the Franco-Cantabrian Glacial Refuge was a Major Source for the European Gene Pool. *Am. J. Hum. Genet.* 75:910-8.

Andrews, R. M., I. Kubacka, P. F. Chinnery et al. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 23:147.

Bolnick, D. A., D. I. Bolnick, and D. G. Smith. 2006. Asymmetric male and female genetic histories among Native Americans from Eastern North America. *Mol. Bio. Evol.* 23:2161-74.

Bortolini, M. C., F. M. Salzano, M. G. Thomas et al. 2003. Y-Chromosome Evidence for Differing Ancient Demographic Histories in the Americas. *Am. J. Hum. Genet*. 73:524–39.

Brown, M. D., S. H. Hosseini, A. Torroni et al. 1998. mtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet*. 63:1852–1861.

Cybulski, J. S. 2010. Human Skeletal Variation and Environmental Diversity in Northwestern North America. In: Auerbach BM, editor. Human Variation in the Americas. Board of Trustees, Southern Illinois University: Center for Archaeological Investigations, 77–112.

Cybulski, J. S. 2006. Skeletal biology: Northwest Coast and Plateau. In: Ubelaker D, editor. Handbook of North American Indians, Vol. 3, Environment, Origins and Population. Washington, DC: Smithsonian Institution, 532–47.

Dulik, M. C., S. I. Zhadanov, L. P. Osipova et al. 2012. Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. *Am. J. Hum. Genet.* 90:229-46.

Elmendorf, W. W. 1965. Linguistic and geographic relations in the Northern Plateau Area. *Southwest J. Anthrop.* 17:365–82.

Eshleman, J. A., R. S. Malhi, J. R. Johnson et al. 2004. Mitochondrial DNA and prehistoric settlements: Native migrations on the western edge of North America. *Hum. Biol.* 76:55-75.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47.

Felsenstein, J. 2005. PHYLIP (phylogeny inference package). Distributed by the author. Seattle (WA): Department of Genome Sciences, University of Washington.

Hünemeier, T., C. E. G. Amorim, S. Azevedo et al. 2012. Evolutionary responses to a constructed niche: ancient Mesoamericans as a model of gene-culture coevolution. *PLoS One* 7:e38862.

Hunley, K., and J. C. Long. 2005. Gene flow across linguistic boundaries in Native North American populations. *Proc. Natl. Acad. Sci. USA* 102:1312-17.

Johnson, J. R., and J. G. Lorenz. 2006. Genetics, Linguistics, and Prehistoric Migrations: An Analysis of California Indian Mitochondrial DNA Lineages. *J. Calif. Gt. Basin. Anthropol.* 26:33–64.

Kaestle, F. A. 1998. Molecular evidence for prehistoric Native American population movement: The Numic expansion. Ph.D. dissertation, University of California, Davis.

Karafet, T. M., F. L. Mendez, M. B. Meilerman et al. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Res.* 18:830-838.

Karafet, T. M., S. L. Zegura, O. Posukh et al. 1999. Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *Am. J. Hum. Genet*. 64:817-31.

Kemp, B. M., A. González-Oliver, R. S. Malhi et al. 2010. Evaluating the farming/language dispersal hypothesis with genetic variation exhibited by populations in the Southwest and Mesoamerica. *Proc. Natl. Acad. Sci. USA* 107:6759-64.

Kemp, B. M., C. Monroe, and D. G. Smith. 2006. Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. *J. Archaeol. Sci*. 33:1680-89.

Lorenz, J. G., and D. G. Smith. 1996. Distribution of four founding mtDNA haplogroups among Native North Americans. *Am. J. Phys. Anthropol*. 101:307–23.

Malhi, R. S., B. M. Kemp, J. A. Eshleman et al. 2007. Mitochondrial haplogroup M discovered in prehistoric North Americans. *J. Archaeol. Sci.* 34:642-48.

Malhi, R. S., K. E. Breece, B. A. Shook, et al. 2004. Patterns of mtDNA diversity in northwestern North America. *Hum. Biol*. 76:33-54.

Miller, S. A., D. D. Dykes, and H. F. Polesky. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16:7164.

Mitchell, D. 1984. Predatory Warfare, Social Status, and the North Pacific Slave Trade. *Ethnology* 23:39-48.

Peakall, R., and P. E. Smouse. 2012 GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537-2539.

Phillips, C., A. Rodriguez, A. Mosquera-Miguel et al. 2008. D9S1120, a simple STR with a common Native American-specific allele: forensic optimization, locus characterization and allele frequency studies. *Forensic Sci. Int.-Gen.* 3:7-13.

Reich, D., N. Patterson, D. Campbell et al. 2012. Reconstructing Native American population history. *Nature* 488:370-4.

Sapir, E., 1915. The Na-Dene languages, a preliminary report. *Am. Anthropol.* 17:534-58.

Schroeder, K. B., T. G. Schurr, J. C. Long et al. 2007. A private allele ubiquitous in the Americas. *Biol. Lett*ers 3:218-23.

Schurr, T. G., M. C. Dulik, A. C. Owings et al. 2012. Clan, language, and migration history has shaped genetic diversity in Haida and Tlingit populations from Southeast Alaska. *Am. J. Phys. Anthropol.* 148:422–35.

Shields, G. F., A. M. Schmiechen, B. L. Frazier et al. 1993. mtDNA Sequences Suggest a Recent Evolutionary Divergence for Beringian and Northern North American Populations. *Am. J. Hum. Genet.* 53:549-62.

Smith, D. G., R. S. Malhi, J. Eshleman et al. 1999. Distribution of mtDNA haplogroup X among native North Americans. *Am. J. Phys. Anthropol*. 110:271-84.

Suttles, W., and W. W. Elmendorf. 1963. Linguistic evidence for Salish prehistory. In Proceedings of the 1962 Annual Spring Meeting of the American Ethnological Society, Symposium on Language and Culture, V. Garfield and W. Chafe, eds. Seattle: University of Washington Press, 42–52.

Tamm, E., T. Kivisild, M. Reidla et al. 2007. Beringian standstill and spread of Native American founders. *PLoS One* 2:e829.

Verdu, P., T. J. Pemberton, R. Laurent et al. 2014. Patterns of admixture and population structure in native populations of Northwest North America. *PLoS Genet.* 10:e1004530.

Wang, S., C. M. Lewis, M. Jakobsson et al. 2007. Genetic variation and population structure in Native Americans. *PLoS Genet.* 3:2049-67.

Ward, R. H., A. Redd, D. Valencia et al. 1993. Genetic and linguistic differentiation in the Americas. *Proc. Natl. Acad. Sci. USA* 90:10663–67.

Zegura, S. L., T. M. Karafet, L. A. Zhivotovsky et al. 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol. Biol. Evol.* 21:164-75.

Zhivotovsky, L. A., N. A. Rosenberg, and M. W. Feldman. 2003. Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. *Am. J. Hum. Genet.* 72:1171–86.

TABLE AND FIGURE CAPTIONS

Table 1. Population sample information. Mitochondrial data is used for all samples in the table. Population samples additionally used for Y Chromosomal and D9S1120 analyses are marked with * and †, respectively, in the Sample Reference column. *N. Am. Interior = North American Interior; N. Am. Coastal = North American Coastal See Figure 1 for a map of geographic region.

Population sample	Study Abbreviation	n	Region*	Language	Language family	Sample Reference
Stswecem'c	ST	30	PNW Inland	Shuswap	Salishan	present study*†
Splatsin	SP	16	PNW Inland	Shuswap	Salishan	present study*†
Laxgalts'ap	LX	20	PNW Coastal	Nisga'a	Tsimshianic	present study*†
Coast Tsimshian	TS	54	PNW Coastal	(Tsimshian)	Tsimshianic	present study*†
Cree	CR	26	N. Am. Interior	Algonquian	Algonquian	present study
Chipewyan	CP	40	N. Am. Interior	Athabaskan	Athabaskan	present study
Dogrib	DG	17	N. Am. Interior	Athabaskan	Athabaskan	Lorenz and Smith 1996; present study
Bella Coola	BC	38	PNW Coastal	Salishan	Salishan	Ward et al. 1993
Haida Gwaii	HG	41	PNW Coastal	Haida	Haida (isolate)	Ward et al. 1993
Kaigani Haida	KH	22	PNW Coastal	Haida	Haida (isolate)	Schurr et al. 2012*
Tlingit (Yakutat and Hoonah)	TL	78	PNW Coastal	Tlingit	Na-Dené	Schurr et al. 2012*
Nuu-chah-nulth	NN	63	PNW Coastal	Nuu-chah-nulth	Wakashan	Ward et al. 1993
Northern Paiute	NP	20	N. Am. Interior	Northern Paiute	Uto-Aztecan	Kaestle, 1998
Yakama	YA	14	N. Am. Interior	Sahaptin	Penutian	Shields et al. 1993
Chumash	CH	16	N. Am. Coastal	(Chumashan)	Chumashan	Johnson and Lorenz 2006
Sisseton/Wahpeton	SW	15	N. Am. Interior	Eastern Dakota	Siouan	Malhi et al. 2001; Bolnick et al, 2006
Aikimel O'odham	AO	56	N. Am. Interior	O/odham	Uto-Aztecan	Kemp et al. 2010
Zuni	ZU	36	N. Am. Interior	Zuni	Zuni isolate	Kemp et al. 2010

Table 2. Mitochondrial haplogroup (A, B, C, D, X, H) frequencies for population samples analyzed in the present study.

Sample	N	A	B	C	D	X	H
Stswecem'c	30	0.33	0.03	0.33	0.27	0.03	0.00
Splatsin	19	0.16	0.05	0.32	0.32	0.11	0.05
Laxgalts'ap	20	0.95	0.00	0.00	0.00	0.05	0.00
Coast Tsimshian	39	0.76	0.02	0.05	0.00	0.17	0.00
Cree	26	0.46	0.00	0.38	0.04	0.12	0.00
Chipewyan	40	1.00	0.00	0.00	0.00	0.00	0.00
Dogrib	17	1.00	0.00	0.00	0.00	0.00	0.00

Table 3. Pairwise F_{ST} matrix generated from mitochondrial sequences for PNW sample populations AO=Akimel O'odham, BC=Bella Coola, CH=Chumash, CP=Chipewyan, CR=Cree, DG=Dogrib, HG=Haida Gwaii, KH=Kaigani Haida, LX= Laxgalts'ap, NN= Nuu-chah-nulth, NP=Northern Paiute, SP=Splatsin, ST= Stswecem'c, SW= Sisseton/Wahpeton, TL= Tlingit (Yakutat and Hoonah), TS=Coast Tsimshian, YA=Yakama, ZU=Zulu

	TL	DG	CP	CR	HG	KH	NN	ST	YA	NP	CH	LX	TS	SP	BC	SW	AO	ZU
TL	0.00																	
DG	0.06	0.00																
CP	0.10	0.10	0.00															
CR	0.28	0.28	0.41	0.00														
HG	0.11	0.09	0.10	0.24	0.00													
KH	0.04	0.04	0.06	0.19	0.00	0.00												
NN	0.22	0.23	0.28	0.03	0.17	0.14	0.00											
ST	0.32	0.34	0.46	0.00	0.28	0.24	0.03	0.00										
YA	0.42	0.46	0.63	0.13	0.43	0.34	0.14	0.12	0.00									
NP	0.49	0.53	0.67	0.16	0.49	0.42	0.17	0.13	0.09	0.00								
CH	0.21	0.26	0.37	0.14	0.17	0.13	0.10	0.14	0.25	0.30	0.00							
LX	0.04	0.18	0.22	0.27	0.09	0.03	0.20	0.31	0.41	0.48	0.20	0.00						
TS	0.09	0.13	0.19	0.14	0.10	0.06	0.10	0.18	0.26	0.33	0.11	0.05	0.00					
SP	0.45	0.52	0.67	0.06	0.43	0.38	0.09	0.02	0.15	0.08	0.23	0.47	0.30	0.00				
BC	0.12	0.17	0.23	0.10	0.10	0.07	0.06	0.11	0.21	0.27	0.09	0.08	0.05	0.21	0.00			
SW	0.26	0.31	0.46	0.05	0.25	0.18	0.05	0.06	0.08	0.14	0.10	0.22	0.10	0.11	0.09	0.00		
AO	0.47	0.46	0.54	0.20	0.44	0.39	0.23	0.20	0.10	0.08	0.34	0.44	0.32	0.20	0.30	0.20	0.00	
ZU	0.52	0.54	0.64	0.34	0.52	0.46	0.33	0.34	0.20	0.23	0.43	0.51	0.37	0.38	0.37	0.32	0.08	0.00

Table 4. Y-chromosomal haplotype frequencies for males in population samples analyzed in the present study.

Sample	N	C/C3	Q/Q3	R1b	I1	I2b1	J1	L	G2a
Stswecem'c	10	0.08	0.33	0.58	0.00	0.00	0.00	0.00	0.00
Splatsin	11	0.18	0.09	0.55	0.00	0.00	0.09	0.09	0.00
Laxgalts'ap	9	0.11	0.44	0.33	0.00	0.11	0.00	0.00	0.00
Coast Tsimshian	27	0.04	0.33	0.37	0.15	0.07	0.00	0.00	0.04

Table 5. D9S1120 9RA frequencies for the PNW study samples.

Sample	N	9	10	11	12	13	14	15	16	17	18	19
Stswecem'c	31	0.15	0.00	0.00	0.23	0.05	0.00	0.13	0.34	0.02	0.08	0.02
Splatsin	15	0.03	0.00	0.03	0.43	0.00	0.00	0.07	0.17	0.10	0.17	0.00
Laxgalts'ap	19	0.18	0.00	0.00	0.03	0.00	0.00	0.21	0.21	0.32	0.05	0.00
Coast Tsimshian	54	0.14	0.00	0.01	0.11	0.02	0.01	0.13	0.36	0.19	0.03	0.00

Figure 1. Map depicting the approximate locations of study populations. See Table 1 for sample information. AO=Akimel O'odham, BC=Bella Coola, CH=Chumash, CP=Chipewyan, CR=Cree, DG=Dogrib, HG=Haida Gwaii, KH=Kaigani Haida, LX= Laxgalts'ap , NN= Nuu-chah-nulth, NP=Northern Paiute, SP=Splatsin, ST= Stswecem'c, SW= Sisseton/Wahpeton, TL= Tlingit (Yakutat and Hoonah), TS=Coast Tsimshian, YA=Yakama, ZU=Zulu

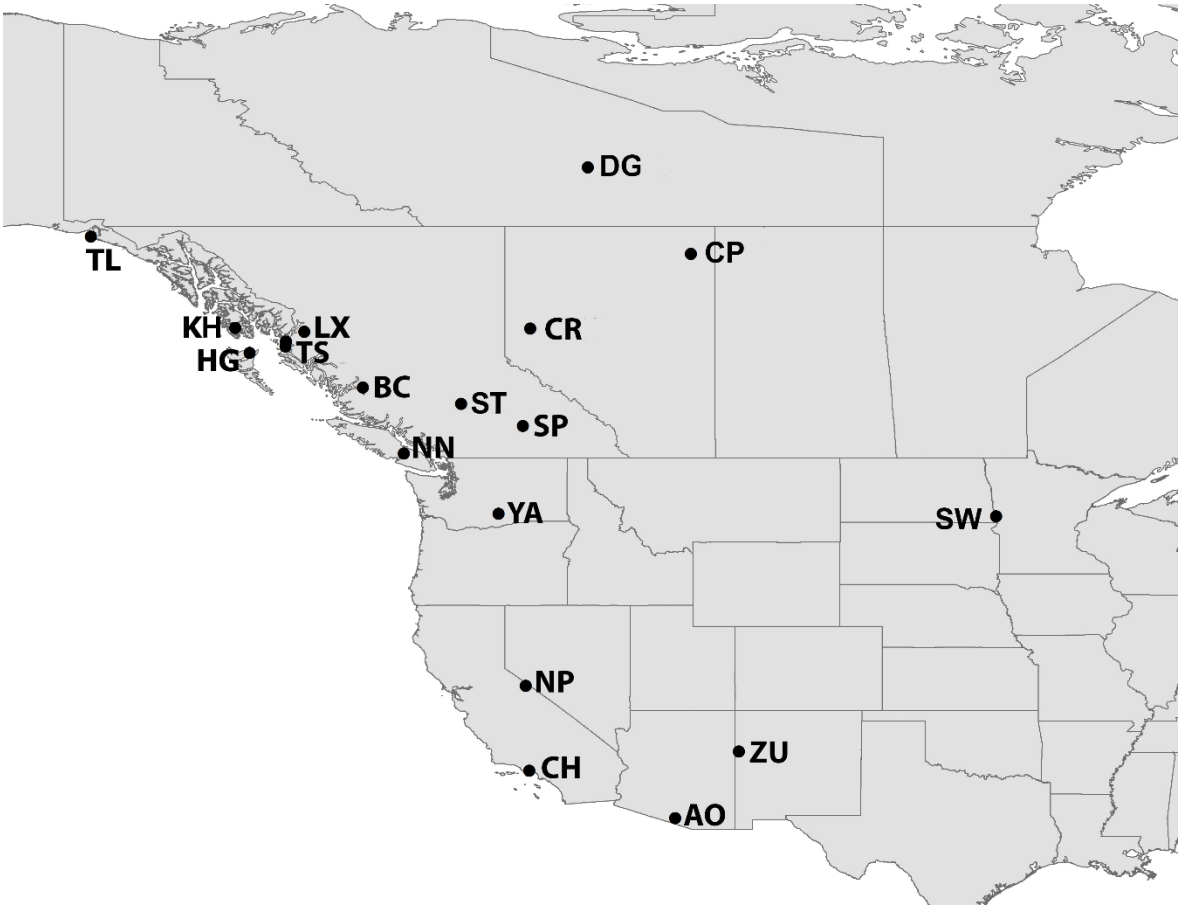

Figure 2. Principal coordinates plot of the pairwise F_{ST} matrix generated from mtDNA sequences for PNW sample populations. See Table 1 for sample information. AO=Akimel O'odham, BC=Bella Coola, CH=Chumash, CP=Chipewyan, CR=Cree, DG=Dogrib, HG=Haida Gwaii, KH=Kaigani Haida, LX= Laxgalts'ap, NN= Nuu-chah-nulth, NP=Northern Paiute, SP=Splatsin, ST= Stswecem'c, SW= Sisseton/Wahpeton, TL= Tlingit (Yakutat and Hoonah), TS=Coast Tsimshian, YA=Yakama, ZU=Zulu. Color indicates geographic affiliations in Table 1: red = North American interior, green = North American coastal, teal = PNW coastal, purple = PNW inland.

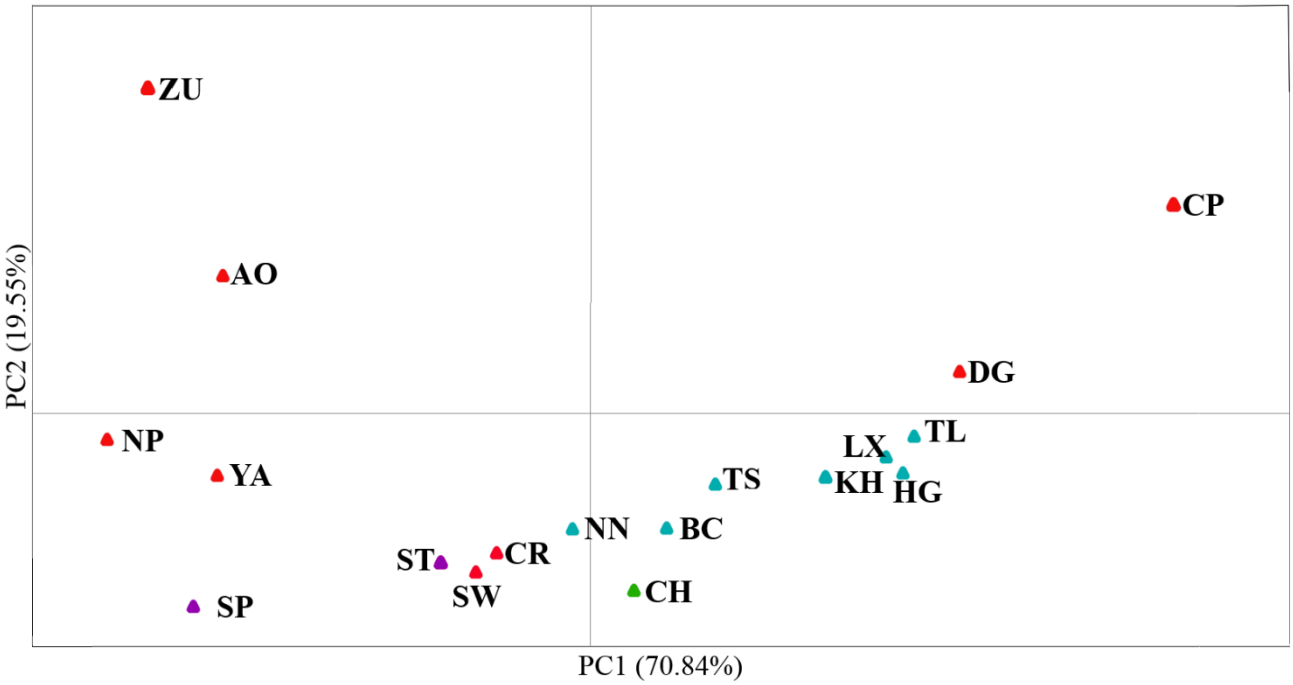

Figure 3. Fitch-Margoliash Least Squares derived unrooted tree representing the pairwise F_{ST} matrix generated from mitochondrial sequences for PNW sample populations and select northern North American samples. BC=Bella Coola, DG=Dogrib, HG=Haida Gwaii, KH=Kaigani Haida, LX= Laxgalts'ap, NN= Nuu-chah-nulth, SP=Splatsin, ST= Stswecem'c, TL= Tlingit (Yakutat and Hoonah), TS=Coast Tsimshian, YA=Yakama. Color indicates geographic affiliations in Table 1: red = North American interior), teal = PNW coastal, purple = PNW inland.

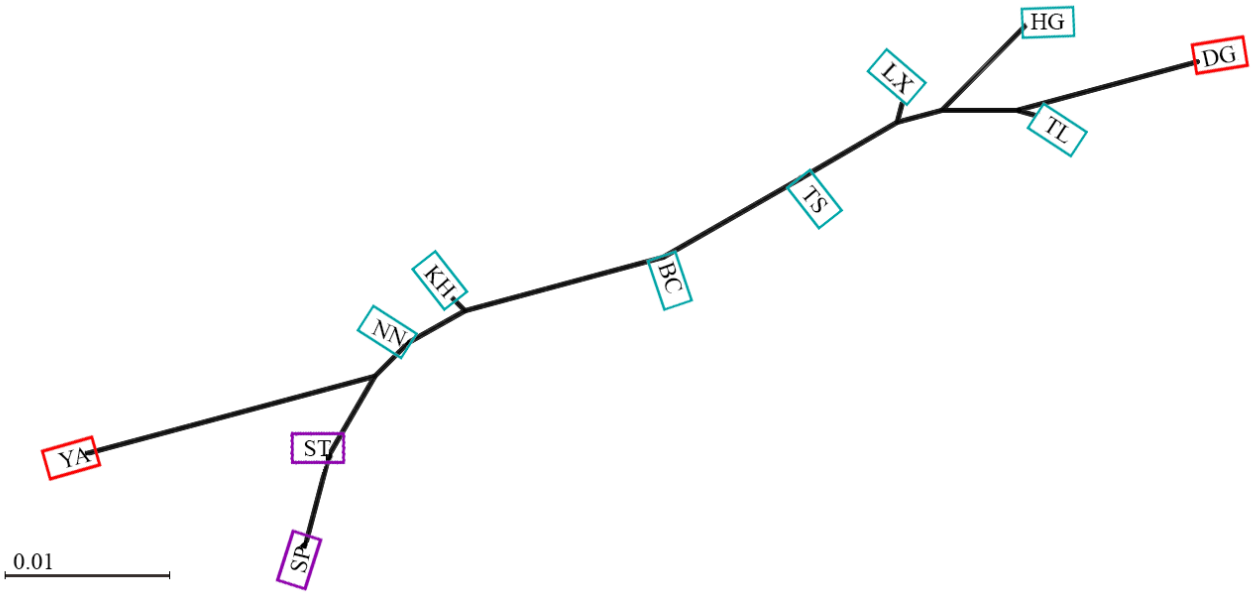

Figure 4. Fitch-Margoliash Least Squares derived unrooted tree representing the Y-STR pairwise F_{ST} matrix for PNW sample populations.

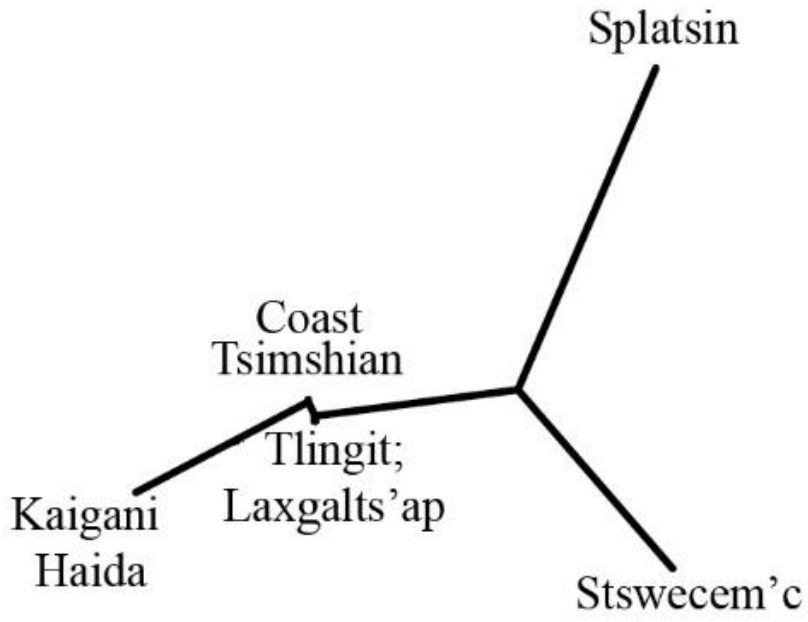

Figure 5. Fitch-Margoliash Least Squares derived unrooted tree representing the pairwise F_{ST} matrix for D9S1120 allelic variation for PNW sample populations.