Regional Information Report 5J18-04

Identification of Genetic Markers Useful for Mixed Stock Analysis of Chinook Salmon in Cook Inlet, Alaska

by

Tyler H. Dann

Chris Habicht

William D. Templin

Lisa W. Seeb

Garrett McKinney

and

James E. Seeb

November 2018

Alaska Department of Fish and Game **Division of Commercial Fisheries**

Symbols and Abbreviations

The following symbols and abbreviations, and others approved for the Système International d'Unités (SI), are used without definition in the following reports by the Divisions of Sport Fish and of Commercial Fisheries: Fishery Manuscripts, Fishery Data Series Reports, Fishery Management Reports, and Special Publications. All others, including deviations from definitions listed below, are noted in the text at first mention, as well as in the titles or footnotes of tables, and in figure or figure captions.

Weights and measures (metric)

centimeter	cm
deciliter	dL
gram	g
hectare	ha
kilogram	kg
kilometer	km
liter	L
meter	m
milliliter	mL
millimeter	mm

Weights and measures (English)

cubic feet per second	ft^3/s
foot	ft
gallon	gal
inch	in
mile	mi
nautical mile	nmi
ounce	oz
pound	lb
quart	qt
yard	yd

Time and temperature

day	d
degrees Celsius	°C
degrees Fahrenheit	°F
degrees kelvin	K
hour	h
minute	min
second	s

Physics and chemistry

all atomic symbols	
alternating current	AC
ampere	A
calorie	cal
direct current	DC
hertz	Hz
horsepower	hp
hydrogen ion activity (negative log of)	pH
parts per million	ppm
parts per thousand	ppt, ‰
volts	V
watts	W

General

Alaska Administrative Code	AAC
all commonly accepted abbreviations	e.g., Mr., Mrs., AM, PM, etc.
all commonly accepted professional titles	e.g., Dr., Ph.D., R.N., etc.
at	@
compass directions:	
east	E
north	N
south	S
west	W
copyright	©
corporate suffixes:	
Company	Co.
Corporation	Corp.
Incorporated	Inc.
Limited	Ltd.
District of Columbia	D.C.
et alii (and others)	et al.
et cetera (and so forth)	etc.
exempli gratia (for example)	e.g.
Federal Information Code	FIC
id est (that is)	i.e.
latitude or longitude	lat or long
monetary symbols (U.S.)	$, ¢
months (tables and figures): first three letters	Jan,...,Dec
registered trademark	®
trademark	™
United States (adjective)	U.S.
United States of America (noun)	USA
U.S.C.	United States Code
U.S. state	use two-letter abbreviations (e.g., AK, WA)

Mathematics, statistics

all standard mathematical signs, symbols and abbreviations	
alternate hypothesis	H_A
base of natural logarithm	*e*
catch per unit effort	CPUE
coefficient of variation	CV
common test statistics	(F, t, χ^2, etc.)
confidence interval	CI
correlation coefficient (multiple)	R
correlation coefficient (simple)	r
covariance	cov
degree (angular)	°
degrees of freedom	df
expected value	*E*
greater than	>
greater than or equal to	≥
harvest per unit effort	HPUE
less than	<
less than or equal to	≤
logarithm (natural)	ln
logarithm (base 10)	log
logarithm (specify base)	log_2, etc.
minute (angular)	'
not significant	NS
null hypothesis	H_O
percent	%
probability	P
probability of a type I error (rejection of the null hypothesis when true)	α
probability of a type II error (acceptance of the null hypothesis when false)	β
second (angular)	"
standard deviation	SD
standard error	SE
variance	
population	Var
sample	var

REGIONAL INFORMATION REPORT 5J18-04

IDENTIFICATION OF GENETIC MARKERS USEFUL FOR MIXED STOCK ANALYSIS OF CHINOOK SALMON IN COOK INLET, ALASKA

by

Tyler H. Dann
Christopher Habicht
William D. Templin
Alaska Department of Fish and Game, Division of Commercial Fisheries, Anchorage

and

Lisa W. Seeb
Garrett McKinney
James E. Seeb
University of Washington, School of Aquatic and Fishery Sciences, Seattle. WA 98105

Alaska Department of Fish and Game
Division of Commercial Fisheries
333 Raspberry Road, Anchorage, Alaska, 99518-1565

November 2018

This research was partially funded by Alaska Sustainable Salmon Fund award 44914 to the University of Washington from the National Oceanic and Atmospheric Administration, US Department of Commerce, administered by the Alaska Department of Fish and Game.

The Regional Information Report Series was established in 1987 and was redefined in 2007 to meet the Division of Commercial Fisheries regional need for publishing and archiving information such as area management plans, budgetary information, staff comments and opinions to Alaska Board of Fisheries proposals, interim or preliminary data and grant agency reports, special meeting or minor workshop results and other regional information not generally reported elsewhere. Reports in this series may contain raw data and preliminary results. Reports in this series receive varying degrees of regional, biometric and editorial review; information in this series may be subsequently finalized and published in a different department reporting series or in the formal literature. Please contact the author or the Division of Commercial Fisheries if in doubt of the level of review or preliminary nature of the data reported. Regional Information Reports are available through the Alaska State Library and on the Internet at: http://www.adfg.alaska.gov/sf/publications/.

Note: Product names used in this publication are included for completeness and do not constitute product endorsement. The Alaska Department of Fish and Game does not endorse or recommend any specific company or their products.

Tyler H. Dann, Christopher Habicht, and William D. Templin,
Alaska Department of Fish and Game, Division of Commercial Fisheries,
333 Raspberry Road, Anchorage, AK 99518, USA

and

Lisa W. Seeb, Garrett McKinney, and James E. Seeb,
University of Washington, School of Aquatic and Fishery Sciences, Seattle. WA 98105

This document should be cited as follows:
Dann, T. H., C. Habicht, W. D. Templin, L. W. Seeb, G. McKinney, and J. E. Seeb. 2018. Identification of genetic markers useful for mixed stock analysis of Chinook salmon in Cook Inlet, Alaska. Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report 5J18-04, Anchorage.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF APPENDICES

ABSTRACT

Recent declines in abundances of Chinook salmon *Oncorhynchus tshawytscha* populations in Cook Inlet, Alaska, have led to fishery restrictions and designation of some populations as *stocks of concern* by the Alaska Board of Fisheries. The utility of genetic tools to inform fishery management in the area has been limited by the lack of genetic differentiation with existing genetic markers among key groups of populations, notably West Cook Inlet and Upper Yentna populations. We applied restriction-site associated DNA sequencing (RAD sequencing) to genotype tens of thousands of novel genetic markers in 283 individuals from 5 populations of Chinook salmon from northwestern Cook Inlet. After quality filtering, 282 individuals were genotyped for 13,804 markers and available for analysis. We selected 500 markers for development of a Cook Inlet panel based upon genetic differentiation among the 5 populations. Overall F_{ST}, a measure of genetic differentiation, averaged 0.003 among all 13,804 markers and 0.029 among the 500 selected for panel development; differentiation among subsets of populations for the 500 panel markers was much higher. Following panel optimization, 183 of the 500 panel markers were incorporated into a panel with previously developed assays. A majority of the markers (171) included in the final panel were located on a linkage map of the Chinook salmon genome. Preliminary baseline investigations suggest that the West Cook Inlet group of populations is now distinguishable from Upper Yentna populations and that Deshka River Chinook salmon are also identifiable from all other Cook Inlet populations. These new markers should be useful in mixed stock analysis to inform management of Cook Inlet fisheries.

Key words: Cook Inlet, Chinook salmon, genetic markers, SNP

INTRODUCTION

The use of mixed stock analysis (MSA) to identify discrete stocks of Pacific salmon (*Oncorhynchus spp.*) has become a key feature of conservation and management activities in Alaska, where real-time assessment of fisheries can maximize harvest of abundant stocks while minimizing risk of overharvesting weaker stocks (Dann et al. 2013). Postseason analyses of mixed stock fisheries are vital to accurately estimate productivities and escapement goals (Cunningham et al. 2017) and inform allocative decisions (Shedd et al. 2016). While broad-scale reporting groups have been successfully identified through MSA, fine-scale stock differentiation is still problematic in many areas due to the low degree of structure exhibited among some populations of Chinook salmon (*O. tshawytscha*; Larson et al. 2014a). Shallow population structure often results in the inability to identify management units in MSA—severely limiting the value of genetic tools for fisheries management. One approach to resolving fine-scale population structure is by selecting panels of hundreds of highly informative markers for MSA from thousands of newly discovered single nucleotide polymorphisms (SNPs). While complete genome assemblies are becoming increasingly available for salmonid species (e.g., Chinook salmon, Christensen et al. 2018; Atlantic salmon, Lien et al. 2016), their utility and cost effectiveness for management applications is less than panels of tens to hundreds of genetic markers.

In recent years, populations of Chinook salmon have been declining in West Cook Inlet, reflecting the general trend across Alaska (Hollowell et al. 2013). During 2008–2016, escapement goals were not achieved for many populations in the area including Alexander Creek (9 of 9 years), Chuitna River (5 of 9 years), Lewis River (9 of 9 years), Talachulitna River (3 of 9 years), and Theodore River (9 of 9 years; Munro and Volk 2017). Due to failures to meet escapement goals, severe Chinook salmon restrictions were enacted on commercial fisheries; sport fisheries have been closed since 2011 on the Theodore, Lewis and Chuitna rivers, and the area had greatly reduced subsistence harvests (Shields and Frothingham 2018).

Chinook salmon have historically made up approximately half of the Tyonek subsistence harvest; however, the Tyonek Chinook salmon catch declined by roughly 60% between 2004 and

2012 (Shields and Dupuis 2013). The subsistence fishery has failed to meet amounts necessary for subsistence in recent years (ADF&G 2013), and residents are fishing longer and relying more upon other salmon species to meet their subsistence needs (Jones and Koster 2018). Chinook salmon caught in the Tyonek subsistence fishery are believed to originate from streams in northwestern Cook Inlet and the Susitna/Yentna river system. Major systems in this region (Beluga, Theodore, Lewis, and Chuitna rivers) all have regulatory restrictions on harvests due to recent low returns, and 6 Chinook salmon stocks in the Northern District of Cook Inlet have been designated as *stocks of concern* by the Alaska Board of Fisheries. Compounding these issues is the fact that it is difficult to differentiate populations on the coast of West Cook Inlet from populations in the upper Yentna River using the current genetic baseline (Barclay and Habicht 2015).

The Alaska Department of Fish and Game (ADF&G) recommended a comprehensive program to estimate stock-specific marine harvests of Chinook salmon in mixed stock fisheries of Cook Inlet using MSA and CWT (ADF&G Chinook Salmon Research Team 2013). Additional sampling of West Cook Inlet populations and development of markers that could differentiate among populations or groups of populations (stocks) in the area would improve the resolution beyond the existing MSA baselines in use at that time. Distinguishing upper Yentna River populations (Upper Yentna stock) from the populations on the coast of West Cook Inlet (West Cook Inlet stock) would improve assessment of the effect of subsistence and commercial harvests on the less abundant West Cook Inlet stock. Combining estimates of stock-specific harvests with escapement estimates on the contributing rivers would enable estimates of total run, harvest rates, and more accurate brood tables. Ultimately, improved estimates of total run and productivity would facilitate better escapement goals and further the mission of managing area fisheries on the sustained yield principle. Without the ability to separate these stocks, uncertainty remains regarding the portions of the harvest assigned to the West Cook Inlet stock and the more abundant Upper Yentna stock.

The Alaska Sustainable Salmon Fund funded research to improve ADF&G's ability to separate stocks of Chinook salmon from northwestern Cook Inlet. Specifically, the Alaska Sustainable Salmon Fund Project number 44914 (Enhanced monitoring of metapopulations of Chinook salmon in Western Cook Inlet and Western Alaska using mapped DNA markers), a collaboration between the University of Washington and ADF&G, sought to 1) develop a linkage map (an ordering and spacing of genetic markers along chromosomes) of Chinook salmon that incorporates variation from Alaska populations; 2) identify novel informative markers for MSA in northwestern Cook Inlet; and 3) utilize the linkage map to optimize the development of marker panels for use in Cook Inlet and Western Alaska. This report describes the identification of markers for panel development and subsequent use in MSA in Cook Inlet (Objective 2).

OBJECTIVES

Identify genetic markers that can differentiate populations or population groups of Chinook salmon from northwestern Cook Inlet and the adjacent Susitna/Yentna river system for panel development and subsequent use in MSA.

METHODS

EXPERIMENTAL DESIGN AND SELECTION OF COLLECTIONS FOR ANALYSIS

We selected 5 populations to use in the identification of informative markers for mixed stock analysis (MSA) in Cook Inlet; these populations were chosen based on their production, importance to subsistence fisheries, and shallow genetic structure based upon a set of 39 SNPs in the legacy baseline (Barclay and Habicht 2015). Four populations from West Cook Inlet were selected for analysis (Chuitna River, Coal Creek, Theodore River and Alexander Creek) along with 1 population from the Susitna/Yentna river system (Talachulitna River). The Chuitna River is the southernmost of the populations selected for analysis and enters Cook Inlet near the village of Tyonek. Coal Creek is a tributary of Beluga Lake that drains into Beluga River, the next major tributary north of the Chuitna. Theodore River is the next tributary north of Beluga River, is a low-gradient system and has headwaters near the headwaters of Talachulitna River. Alexander Creek is a low-gradient system that enters Cook Inlet near the mouth of the Susitna River. The Talachulitna River drains a series of tributary lakes and creeks and enters the Skwenta River, a tributary of the Yentna River. Fifty-six and 57 individuals/population were selected for sequencing (Table 1; Figure 1).

COLLECTION AND FILTERING OF SEQUENCE DATA

Sequence Acquisition

We prepared libraries for restriction-site associated DNA (RAD) sequencing following the protocols of Larson et al. (2014b). The restriction enzyme *SbfI* was used to digest the DNA from selected individuals, and *SbfI*-specific Illumina linkers—each containing a unique barcode—were ligated to each digested DNA sample as described in Miller et al. (2012). Individual samples were pooled into 3 libraries each containing 96 individuals. For each sequencing lane, a 12bp (6bp unique barcode plus 6bp *SbfI* site) sequence was used to uniquely identify reads belonging to the multiplexed individuals. Each library was assessed for quality and concentration using a Bioanalyzer DNA 1000 kit (Agilent Technologies, Santa Clara, California), and final concentration was determined by the Bioanalyzer software (Agilent Technologies Clara, California). Ten nanomoles of each library in EB (Qiagen, Valencia, California) and 0.1% Tween 20 were sequenced at the University of Oregon High Throughput Sequencing Facility using the Illumina HiSeq2000 producing 100 bp single-end reads. A second round of sequencing was used to equalize numbers of retained reads among individuals after titrating DNA concentrations.

Stacks Protocol

Bioinformatic analyses including Perl scripts and the Stacks pipeline (Catchen et al. 2013) were used to demultiplex sequence data, call genotypes, and filter loci and individuals for missing data. We used Stacks v1.24 and the following programs and standard parameter values to demultiplex and call genotypes: *process_radtags* (-e sbfI -c -r -q -E phred33 --filter_illumina -i gzfastq -t 94), *ustacks* (-m 2 -M 2 -p 4 -r --model_type bounded --bound_low 0 --bound_high 0.05), *sstacks* (stacks were referenced to a catalog based on linkage map catalog described above after adding 4 individuals from each of the 5 populations), and *populations* (-s -r 0.5 -p 5 -a 0.05 --fstats --genepop --plink –structure). These parameter settings had the effects of removing any reads with uncalled bases or low quality scores and aligning reads with a maximum distance of 2 nucleotides between stacks and minimum depth of coverage of 2. We used the bounded SNP model with an upper error rate bound of 0.05. To ensure that loci were called the same name as used in the linkage map, we built upon the map catalog while incorporating any novel variants from these 5 populations. The *populations* settings ensured that genotypes were scored for only those loci present in all populations, observed in half of all individuals within a population and with a minimum minor allele frequency of 0.05.

Sequence Filtering

We filtered genotype data to remove loci and individuals missing substantial amounts of data. Loci and individuals missing 25% or more of their genotypic data were filtered with the program plink v1.9 (Chang et al. 2015). We quantified the genotyping rate as the percent of the total number of possible individual, locus-specific genotypes that were called at each filtering step. We tested for conformance to Hardy-Weinberg expectations using exact tests in plink and removed loci with p-values < 0.05 in 3 or more populations after applying the suggested mid-p adjustment to improve power with large datasets (Graffelman and Moreno 2013). Lastly, we selected a single SNP from each RAD tag for subsequent analyses. A *genepop* file (Rousset 2008) of final filtered data was written out of plink for subsequent analyses in R (R Core Team 2018).

EXAMINATION OF POPULATION STRUCTURE AND SELECTION OF MARKERS FOR PANEL DEVELOPMENT

Examination of Population Structure

We visualized genetic structure with individual-based principal component analysis (PCA). We examined PCAs for the same individuals that passed the missing data filter using 3 marker sets: (1) the 39 of the original 42 SNPs from the legacy baseline that were variable in Cook Inlet, (2) the entire dataset after filtering, and (3) the subset of SNPs identified for panel development (described below).

Estimates of F_{ST} and Weighting of SNPs for Panel Development

We estimated F_{ST} values based upon Weir and Cockerham (1984) to assess variation among populations for each SNP that passed filters to identify markers for panel development. We estimated F_{ST} values among all populations, between each population and the remaining 4, and among the 3 populations (Chuitna, Alexander and Theodore) that exhibited the least structure and clustered together in PCAs. We ranked each SNP based upon these F_{ST} estimates and selected a panel of 500 to maximize divergence among populations with the following weighting:

(1) 300 SNPs were selected based upon F_{ST} among the Chuitna, Alexander, and Theodore populations.
(2) 100 SNPs were selected based upon F_{ST} between Talachulitna and the remaining 4 populations.
(3) 100 SNPs were selected based upon F_{ST} between Coal and the remaining 4 populations.

Lastly, we located each SNP identified for panel development on the combined Washington-Alaska linkage map (McKinney et al. *In press*) to characterize their distribution on the genome and identify putative genomic hotspots where divergence among Cook Inlet populations is greatest.

Examination of Likelihood Profiles

We also examined likelihood profiles for each population by calculating the self-assignment probability for each individual within populations. The likelihood of each individual's genotype originating from all populations based on leave-one-out population allele frequencies was calculated following the methods of gsi_sim (Anderson et al. 2008). These genotype likelihoods were summarized to the population level to determine the probability of individuals from a population being assigned back to that population. We visualized these probabilities as a matrix to better understand self-assignment of individuals back to their population of origin, and to gain insight into potential misallocation in fishery mixtures for the original 39 SNPs available from the legacy baseline (Barclay and Habicht 2015) and the subset of SNPs identified for panel development. We did not estimate likelihoods based upon the entire, filtered RAD dataset (13,804 loci) due to computational limitations.

Panel Development

We combined the 500 SNPs selected for panel development and additional Taqman assay-based markers developed elsewhere for a final panel that would multiplex together using Genotyping-in-Thousands by sequencing (GTseq) chemistry (Campbell et al. 2015). These additional Taqman assays included 6 SNPs from the existing Cook Inlet baseline (Ots_HGFA-446, Ots_HSP90B-385, Ots_il-1racp-166, Ots_ins-115, Ots_ZNF330-181). Panel development included filtering out loci with (1) SNPs in a required primer region of each RAD amplicon (bases 1-20 and 75-94 of the 94 bp amplicons); (2) amplicons with repeated elements that would amplify multiple regions of the genome with the online software RepeatMasker (Smit et al. 2018); and (3) amplicons that generated excessive sequence to the detriment of remaining amplicons in multiple rounds of panel testing.

RESULTS

COLLECTION AND FILTERING OF SEQUENCE DATA

A total of 376,355,791 retained reads were processed through the Stacks pipeline (Table 1, Figure 2). The 5 populations had similar quality of sequence data, averaging 75M retained reads while individuals averaged 1.3M retained reads. After processing sequence data through the Stacks pipeline, 33,407 loci were genotyped for the 283 individuals at an overall genotyping rate of 72.4% (Table 2). Filtering for missing data at the 25% threshold resulted in 20,133 loci (60.3% of the 33,407 loci); just 1 individual was removed for missing data (99.6% were retained). A total of 3,761 loci were removed for failing to conform to Hardy-Weinberg expectations (resulting in 16,372 retained loci). After filtering for a single SNP/amplicon, 13,804 loci genotyped for 282 individuals were available for subsequent analyses at an overall genotyping rate of 92.9% (Table 2).

EXAMINATION OF POPULATION STRUCTURE AND SELECTION OF MARKERS FOR PANEL DEVELOPMENT

Examination of Population Structure

PCAs based upon the 3 marker sets depicted markedly different patterns of population genetic structure (Figure 3). The legacy set of 39 SNPs previously used in mixed stock analysis indicated no structure among populations; all variation was based upon individual genotypic variability.

The PCA based upon the entire RAD dataset of 13,804 SNPs suggested divergence of Coal Creek and the Talachulitna River from the remaining 3 populations which all clustered closely together. The 500 SNPs selected for panel development achieved the greatest divergence among populations, with the Chuitna River diverging from Theodore River and Alexander Creek, but still little segregation between those 2 drainages.

Estimates of F_{ST} and Results of Panel Development

Estimates of F_{ST} exhibited substantial contrast among loci at differentiating populations. F_{ST} among all populations averaged 0.003 for the entire RAD dataset; 24 loci had an $F_{ST} > 0.05$ (Figure 4). Distributions of overall F_{ST} for the 3 marker sets were markedly different, with the 500 selected for panel development (Appendix 1) exhibiting the greatest divergence among populations (Figure 5). Of the 500 selected for panel development, primers were able to be designed for 183 (37%) that multiplexed well together. Of these final SNPs in the panel, 89 were included to differentiate Chuitna-Theodore-Alexander (mean F_{ST} = 0.048), 44 to differentiate Coal from other populations (mean F_{ST} = 0.073), and 50 to differentiate Talachulitna from other populations (mean F_{ST} = 0.079; Table 3).

A majority of the 500 selected for panel development (461; Appendix 1) and 171 of the 183 loci included in the final panel were located on the linkage map. Loci in the final panel were located on all linkage groups, ranging from 1 to 11 loci/linkage group (mean = 5; Table 4) and broadly distributed throughout and among linkage groups (Figure 6).

Examination of Likelihood Profiles

Likelihood profiles indicated high uncertainty in genetic assignment to population of origin for the 39 current SNPs, and greatly improved assignment ability with the 500 loci selected for panel development. Theodore River and Alexander River exhibited a greater likelihood of assigning to one another than to themselves with the original 39 SNPs (Figure 7).

DISCUSSION

We applied RAD sequencing to discover genetic markers to differentiate 5 closely related populations of Chinook salmon important to commercial, sport, and subsistence fisheries in Cook Inlet. We selected 500 loci from a filtered dataset of 13,804 loci to maximize divergence among these populations and evaluated population structure and assignment likelihood profiles for existing and final RAD marker sets. No structure among populations was apparent based upon the 39 legacy SNPs while the 500 selected for panel development differentiated populations, although Alexander and Theodore rivers still exhibited shallower structure than the other 3 populations. It is likely that gene flow among these shallow-gradient, coastal populations is more likely than with Coal Creek above Beluga Lake or Talachulitna River much farther up the Susitna River system.

We selected loci based upon the fixation index F_{ST} and weighted the numbers of SNPs selected for panel development by the structure we observed among populations to maximize differentiation among populations. Other approaches to selecting genetic markers for MSA have been suggested, notably random forest (i.e., RF; Sylvester et al. 2018) and backwards elimination locus selection (i.e., BELS; Bromaghin 2008). We have examined the utility of RF for marker selection and found it to be sensitive to the size of marker panel chosen for analysis, often resulting in little contrast in estimated marker power for MSA among loci (ADF&G, unpublished data). BELS has been shown to be effective at identifying subsets of markers for MSA but is too computationally costly for large datasets such as ours (ADF&G, unpublished data). We previously observed high concordance between F_{ST} and utility in MSA when selecting a subset of 24 SNPs from a panel of 96 for Bristol Bay sockeye salmon applications (ADF&G, unpublished data). Each markers' utility in MSA measured with the program *BAYES* (Pella and Masuda 2001) used by ADF&G exhibited nearly perfect correlation with F_{ST} in the selection of the 24 SNPs for Bristol Bay applications. Given that context, we used F_{ST} to select loci for panel development and gave the populations exhibiting the least structure the greatest weight in terms of numbers of loci (Storer et al. 2012).

We expect this panel will be a very useful genetic tool for managing salmon fisheries in Cook Inlet and will improve our understanding of Chinook salmon in the region. While baseline evaluation tests incorporating other Cook Inlet populations would provide a more complete assessment of the new marker sets ability for mixed stock analysis, the leave-one-out likelihood profiles provide a measure of the improvement expected from the new marker set. The final panel incorporating markers from this discovery effort has recently been used to improve the baseline for Cook Inlet Chinook salmon, and preliminary baseline evaluation tests suggest improved resolution among stocks in the region. New stocks are now identifiable in MSA within Cook Inlet and all stocks will be estimated more accurately and precisely with the addition of hundreds of genetic markers. The previous Northwest/Yentna is now split into West and Yentna, a new Deshka group is now separate from the Susitna group that previously included it, and the Kenai/Kasilof group is now split into Kenai Tributaries, Kenai Mainstem and Kasilof Mainstem (Figure 8; ADF&G, unpublished data). The ability to identify these stocks will improve stock assessment so that fisheries in the region can be better managed to maximize sustainable yield, and, hopefully, improve returns and remove *stock of concern* status for those populations currently designated. We expect MSA applications of this new baseline to include reanalysis of subsistence harvests in the Tyonek area with improved resolution and development of a genetic mark–recapture estimate of abundance for the Susitna River based upon weir counts and MSA estimates for the Deshka River stock.

ACKNOWLEDGEMENTS

We thank Carita Pascal, University of Washington, for exceptional laboratory assistance and Andy Barclay, ADF&G, for help selecting populations to analyze, and Heather Hoyt, ADF&G, for facilitating sample sharing, and the Alaska Sustainable Salmon Fund for funding this project. This research was partially funded by Alaska Sustainable Salmon Fund award 44914 to the University of Washington from the National Oceanic and Atmospheric Administration, US Department of Commerce, administered by the Alaska Department of Fish and Game. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the US Department of Commerce, National Oceanic and Atmospheric Administration, or the Alaska Department of Fish and Game.

REFERENCES CITED

ADF&G Chinook Salmon Research Team. 2013. Chinook salmon stock assessment and research plan, 2013. Alaska Department of Fish and Game, Special Publication No. 13-01, Anchorage.

Allendorf, F. W. 2017. Genetics and the conservation of natural populations: allozymes to genomes. Molecular Ecology 26(2):420–430.

Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for predicting the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65:1475–1486.

Barclay, A. W., and C. Habicht. 2015. Genetic baseline for Upper Cook Inlet Chinook salmon: 42 SNPs and 7,917 fish. Alaska Department of Fish and Game, Fishery Manuscript Series No. 15-01, Anchorage.

Bromaghin, J. F. 2008. BELS: backward elimination locus selection for studies of mixture composition or individual assignment. Molecular Ecology Resources 8(3):568–571.

Campbell, N. R., S. A. Harmon, and S. R. Narum. 2015. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15(4):855–867.

Catchen, J., P. Hohenlohe, S. Bassham, A. Amores, and W. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22(11):3124–3140.

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4(1):7.

Christensen, K. A., J. S. Leong, D. Sakhrani, C. A. Biagi, D. R. Minkley, R. E. Withler, E. B. Rondeau, B. F. Koop, and R. H. Devlin. 2018. Chinook salmon (*Oncorhynchus tshawytscha*) genome and transcriptome. PLOS ONE 13(4):e0195461.

Cunningham, C. J., T. A. Branch, T. H. Dann, M. Smith, J. E. Seeb, L. W. Seeb, and R. Hilborn. 2017. A general model for salmon run reconstruction that accounts for interception and differences in availability to harvest. Canadian Journal of Fisheries and Aquatic Sciences 75(3):439–451.

Dann, T. H., C. Habicht, T. T. Baker, and J. E. Seeb. 2013. Exploiting genetic diversity to balance conservation and harvest of migratory salmon. Canadian Journal of Fisheries and Aquatic Sciences 70(5):785–793.

Graffelman, J., and V. Moreno. 2013. The mid p-value in exact tests for Hardy-Weinberg equilibrium. Statistical applications in genetics and molecular biology 12(4):433–448.

Hollowell, G., E. O. Otis, and E. Ford. 2013. 2013 Lower Cook Inlet area finfish management report. Alaska Department of Fish and Game, Fishery Management Report No. 14-30, Anchorage.

Jones, B. E., and D. Koster. 2018. Subsistence harvest and uses of salmon in Tyonek, 2015 and 2016. Alaska Department of Fish and Game, Division of Subsistence, Technical Paper No. 439, Anchorage.

Larson, W. A., J. E. Seeb, C. E. Pascal, W. D. Templin, L. W. Seeb, and E. Taylor. 2014a. Single-nucleotide polymorphisms (SNPs) identified through genotyping-by-sequencing improve genetic stock identification of Chinook salmon (*Oncorhynchus tshawytscha*) from western Alaska. Canadian Journal of Fisheries and Aquatic Sciences 71(5):698–708.

Larson, W. A., L. W. Seeb, M. V. Everett, R. K. Waples, W. D. Templin, and J. E. Seeb. 2014b. Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). Evolutionary Applications 7(3):355–369.

Lien, S., B. F. Koop, S. R. Sandve, J. R. Miller, M. P. Kent, T. Nome, T. R. Hvidsten, J. S. Leong, D. R. Minkley, A. Zimin, F. Grammes, H. Grove, A. Gjuvsland, B. Walenz, R. A. Hermansen, K. von Schalburg, E. B. Rondeau, A. Di Genova, J.K.A. Samy, J. Olav Vik, M. D. Vigeland, L. Caler, U. Grimholt, S. Jentoft, D. Inge Våge, P. de Jong, T. Moen, M. Baranski, Y. Palti, D. R. Smith, J. A. Yorke, A. J. Nederbragt, A. Tooming-Klunderud, K. S. Jakobsen, X. Jiang, D. Fan, Y. Hu, D. A. Liberles, R. Vidal, P. Iturra, S. J. M. Jones, I. Jonassen, A. Maass, S. W. Omholt, and W. S. Davidson. 2016. The Atlantic salmon genome provides insights into rediploidization. Nature 533:200–205.

REFERENCES CITED (Continued)

McKinney, G. J., L. W. Seeb, C. Pascal, W. D. Templin, S. E. Gilk-Baumer, T. H. Dann, and J. E. Seeb. *In press*. Dense SNP panels resolve closely related members of Chinook salmon metapopulations. Evolutionary Applications.

Miller, M. R., J. P. Brunelli, P. A. Wheeler, S. Liu, C. E. Rexroad III, Y. Palti, C.Q. Doe, and G. H. Thorgaard. 2012. A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. Molecular Ecology 21(2):237–249.

Munro, A. R., and E. C. Volk. 2017. Summary of Pacific salmon escapement goals in Alaska with a review of escapements from 2008 to 2016. Alaska Department of Fish and Game, Fishery Manuscript Series No. 17-05, Anchorage.

Pella, J., and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin 99(1):151–167.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rousset, F. 2008. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8(1):103–106.

Shedd, K. R., M. B. Foster, T. H. Dann, H. A. Hoyt, M. L. Wattum, and C. Habicht. 2016. Genetic stock composition of the commercial harvest of sockeye salmon in Kodiak management area, 2014–2016. Alaska Department of Fish and Game, Fishery Manuscript Series No. 16-10, Anchorage.

Shields, P., and A. Dupuis. 2013. Upper Cook Inlet commercial fisheries annual management report. 2013. Alaska Department of Fish and Game, Fishery Management Report No. 13-49, Anchorage.

Shields, P., and A. Frothingham. 2018. Upper Cook Inlet commercial fisheries annual management report, 2017. Alaska Department of Fish and Game, Fishery Management Report No. 18-10, Anchorage.

Smit, A. F. A., R. Hubley, and P. Green. 2018. RepeatMasker Open-4.0. 2013-2015 <http://www.repeatmasker.org>.

Storer, C. G., C. E. Pascal, S. B. Roberts, W. D. Templin, L. W. Seeb, and J. E. Seeb. 2012. Rank and order: Evaluating the performance of SNPs for individual assignment in a non-model organism. PLoS ONE 7(11):e49018.

Sylvester, E. V. A., P. Bentzen, I. R. Bradbury, M. Clément, J. Pearce, J. Horne, and R. G. Beiko. 2018. Applications of random forest feature selection for fine-scale genetic population assignment. Evolutionary Applications 11(2):153–165.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.

TABLES AND FIGURES

Table 1.–Summary of populations of Chinook salmon from Northwest Cook Inlet sequenced to discover markers useful for mixed stock analysis including ADF&G collection code, latitude, longitude, collection date(s), numbers of individuals sequenced (*n*), total retained reads, range of retained reads among individuals, and mean, standard deviation (SD) and coefficient of variation (CV) of reads/individual.

Population	ADF&G code(s)	Lat	Long	Collection Date(s)	*n*	Retained reads	Range	Reads/Individual	SD	CV (%)
Chuitna	KCHUI09	61.14	–151.44	7/14/2009	57	73,568,982	643,684–1,757,213	1,290,684	251,940	19.5
Coal	KCOAL10/11	61.58	–151.77	7/21/2010 and 7/19/2011	56	74,250,630	847,358–1,834,086	1,325,904	223,883	16.9
Theodore	KTHEO12	61.44	–151.09	6/4/2012	56	72,798,733	764,365–1,847,324	1,299,977	225,188	17.3
Alexander	KSUCKCR11/12	61.60	–150.77	7/21/2011 and 7/18/2012	57	77,099,524	756,514–1,995,350	1,352,623	253,193	18.7
Talachulitna	KTALA08/10	61.69	–151.49	7/28/2008 and 7/13/2010	57	78,637,922	828,762–1,850,292	1,379,613	251,936	18.3
Total					283	376,355,791	643,684–1,995,350	1,329,879	242,283	18.2

Table 2.–Numbers of loci and individuals (and percent of starting total) remaining after 4 stages of filtering: initial data from stacks pipeline, after filtering for 25% missing data by loci and then individuals, after removing loci that failed to conform to HWE expectations, and after selecting a single SNP from each RAD tag. Also included is the overall genotyping rate at each stage.

	Loci		Individuals		
Stage	Number	Percent	Number	Percent	Genotyping rate
Initial from stacks	33,407	100	283	100	72.4
Missing data	20,133	60.3	282	99.6	92.4
HWE expectations	16,372	49.0	282	99.6	92.7
Single SNP/RAD tag	13,804	41.3	282	99.6	92.9

Table 3.–Number of amplicons from the 500 selected for panel development that were successfully incorporated into the panel and their mean F_{ST} by purpose. The 3 purposes were to split apart the cluster of Chuitna and Theodore rivers and Alexander Creek (Cluster), differentiate Coal Creek from other populations (Coal) and to differentiate Talachulitna River from other populations (Tala).

Purpose	# amplicons	Mean F_{ST} For Purpose
Cluster	89	0.048
Coal	44	0.073
Tala	50	0.079
Total/Average	183	0.067

Table 4.–Number of amplicons from the 500 selected for panel development that were successfully incorporated into the panel and are located on the linkage map by linkage group.

Linkage group	No. amplicons from selected 500
Ots01	8
Ots02	4
Ots03	5
Ots04	5
Ots05	4
Ots06	4
Ots07	7
Ots08	5
Ots09	7
Ots10	8
Ots11	8
Ots12	11
Ots13	3
Ots14	8
Ots15	7
Ots16	3
Ots17	1
Ots18	2
Ots19	5
Ots20	5
Ots21	3
Ots22	4
Ots23	2
Ots24	2
Ots25	7
Ots26	5
Ots27	3
Ots28	8
Ots29	6
Ots30	4
Ots31	7
Ots32	2
Ots33	6
Ots34	2
Mean	5
Total	171

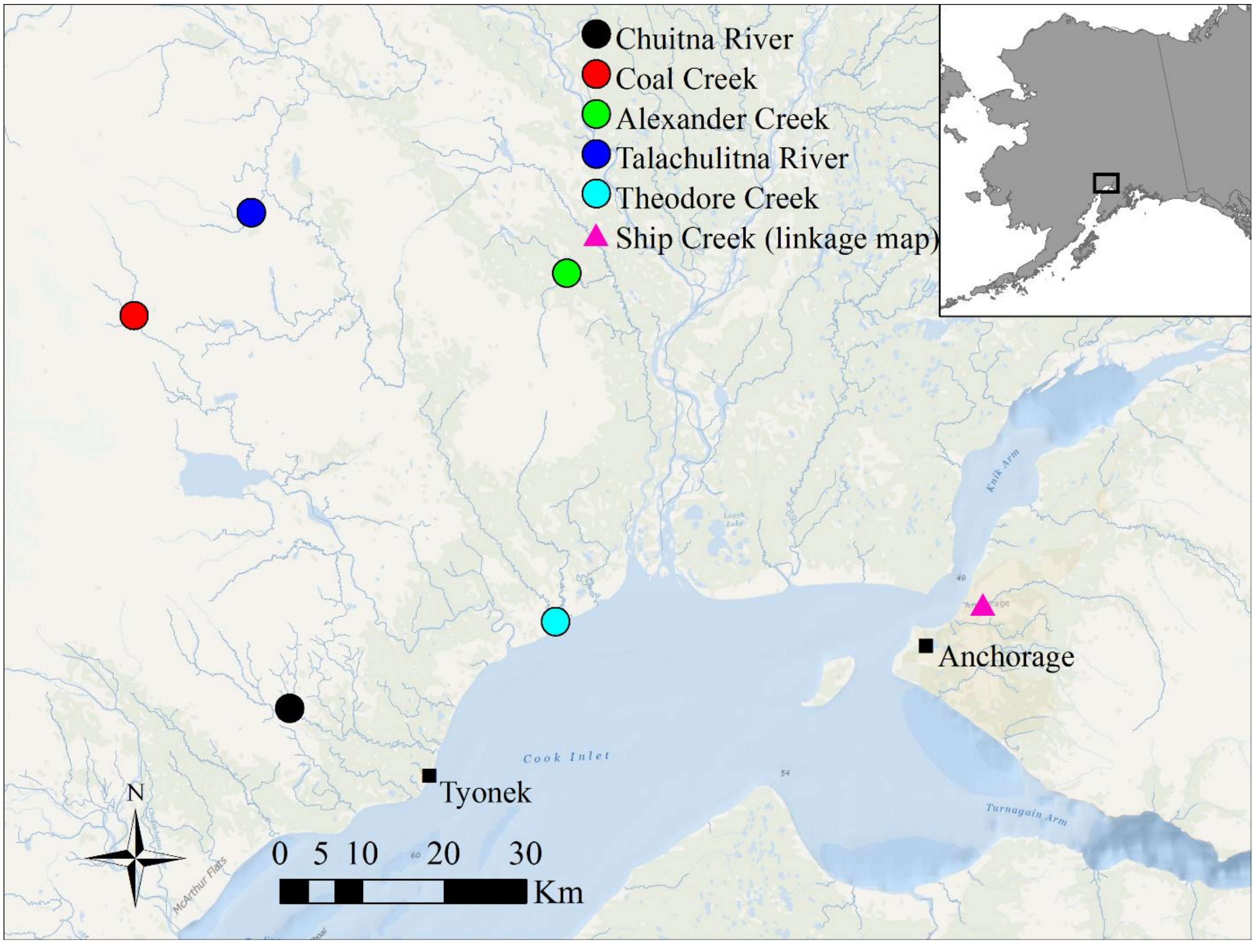

Figure 1.–The location of 5 populations of Chinook salmon used to identify markers for panel development for MSA in Cook Inlet (circles). The families used to construct the combined Washington-Alaska linkage map originated from Ship Creek (triangle).

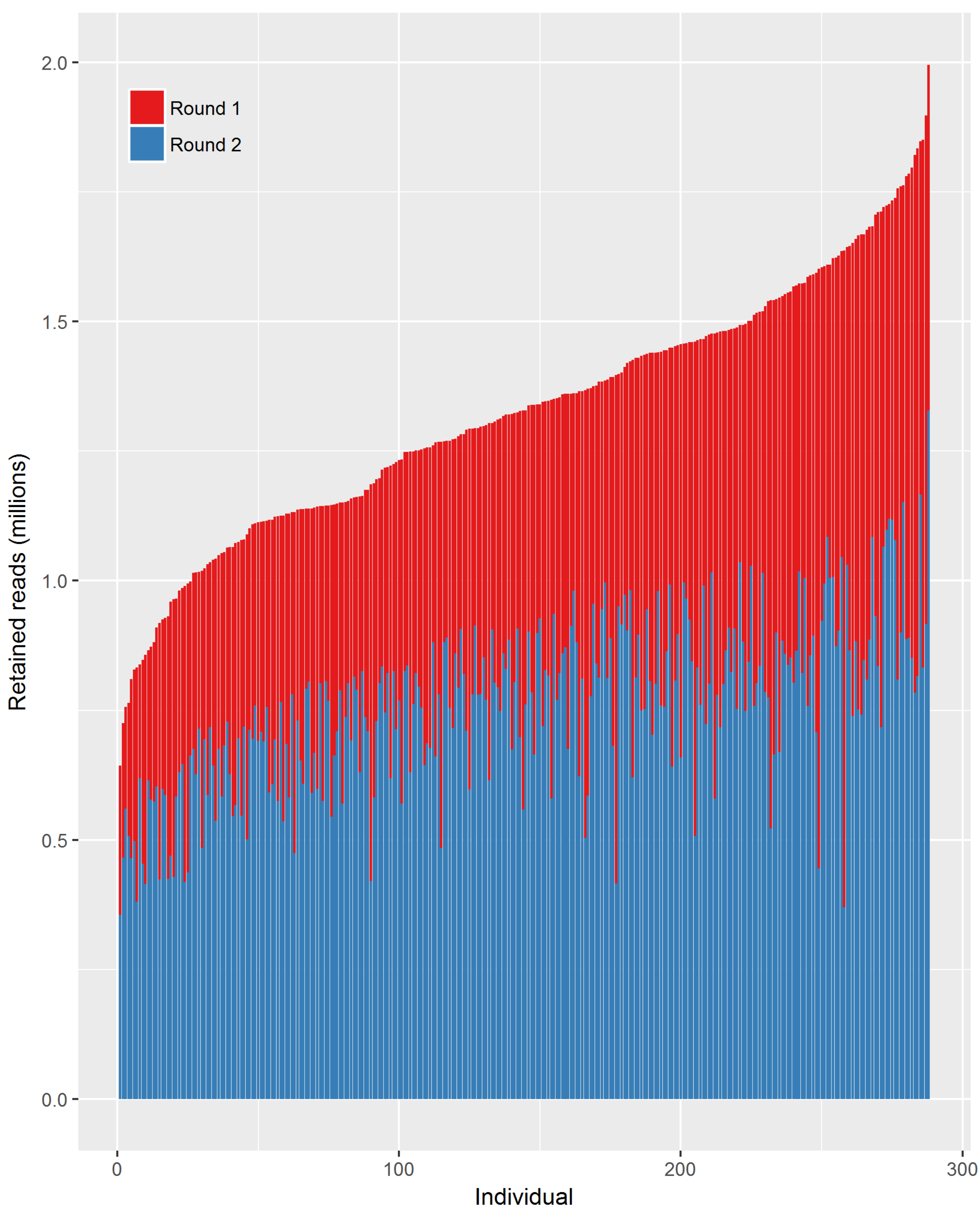

Figure 2.–Number of retained reads by round of sequencing for individuals from 5 Cook Inlet populations of Chinook salmon used to identify markers for panel development for MSA in Cook Inlet.

Figure 3.–Individual-based principal coordinate analysis of 282 individuals from 5 populations of Cook Inlet Chinook salmon used to identify markers for panel development for MSA in Cook Inlet by marker set. Marker sets include the 39 legacy SNPs in the existing baseline (top), the 13,804 RAD SNPs of the final filtered dataset (middle), and 500 SNPs selected for panel development based upon F_{ST} (bottom).

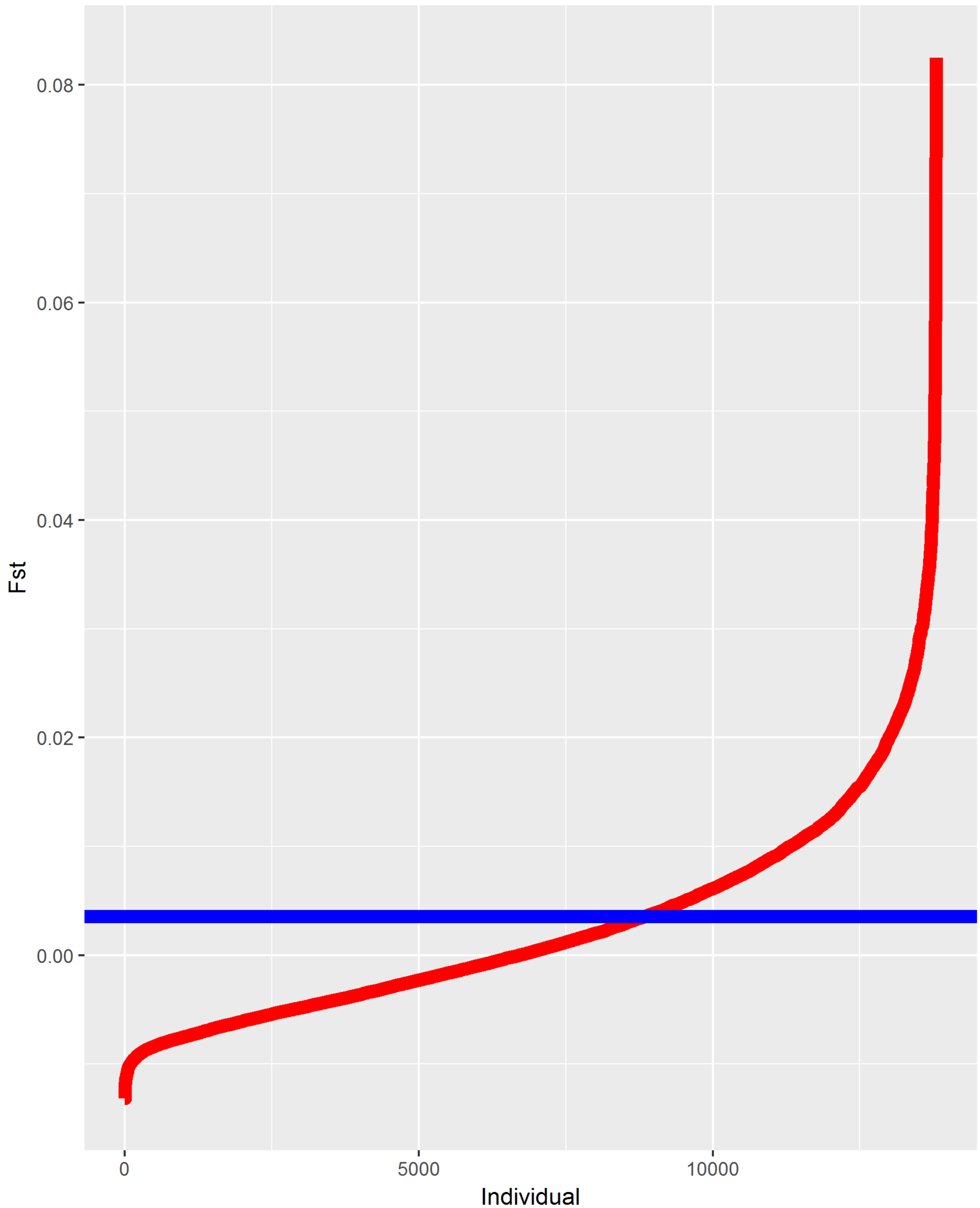

Figure 4.– F_{ST} estimated among all populations (red line) for 13,804 SNPs that passed filters and average F_{ST} among all SNPs (blue line, 0.003).

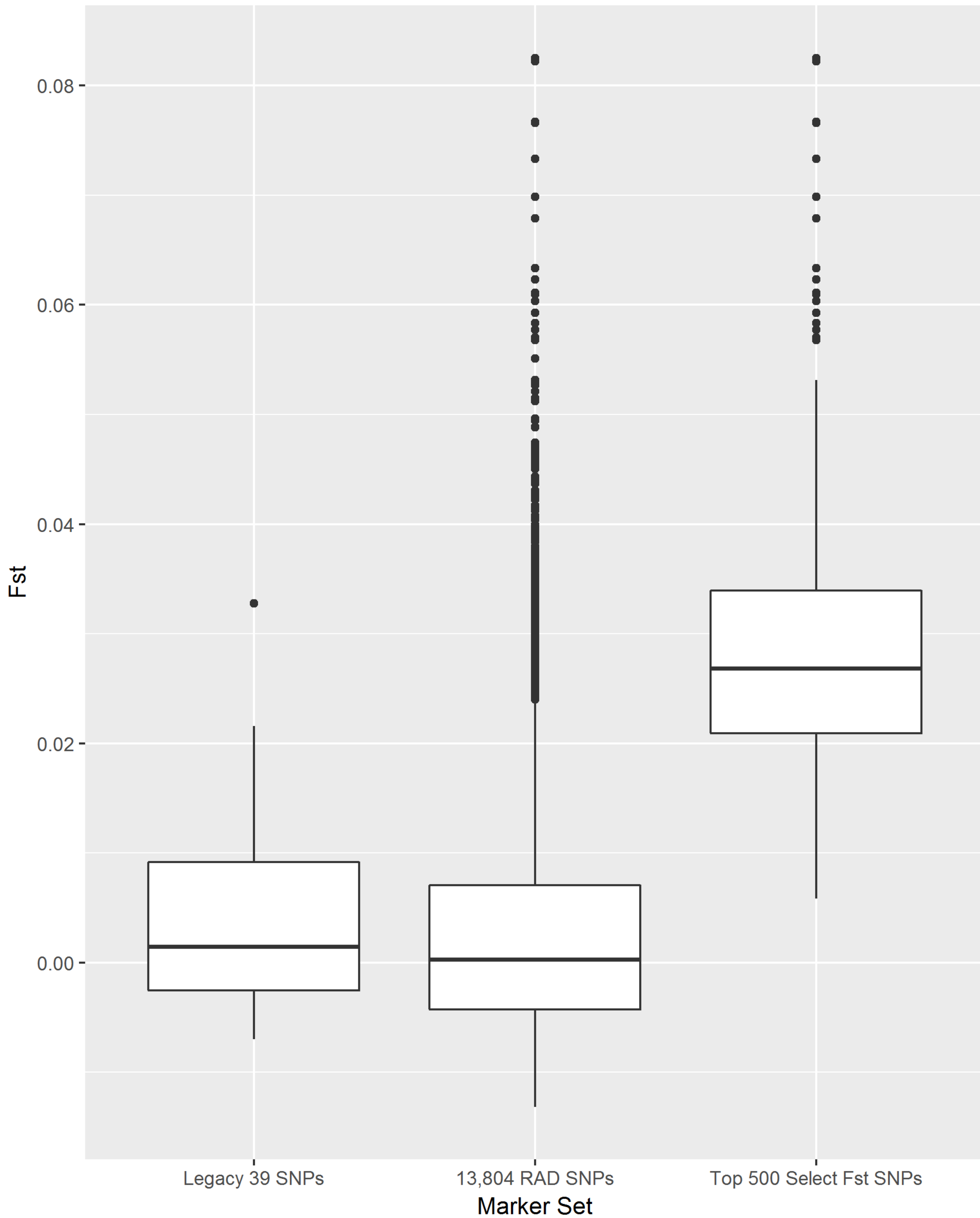

Figure 5.–Distribution of overall F_{ST} among 5 populations of Cook Inlet Chinook salmon used to identify markers for panel development for MSA in Cook Inlet by marker set. Marker sets include the 39 legacy SNPs common to the existing baseline, the 13,804 RAD SNPs of the final filtered dataset and 500 SNPs selected for panel development.

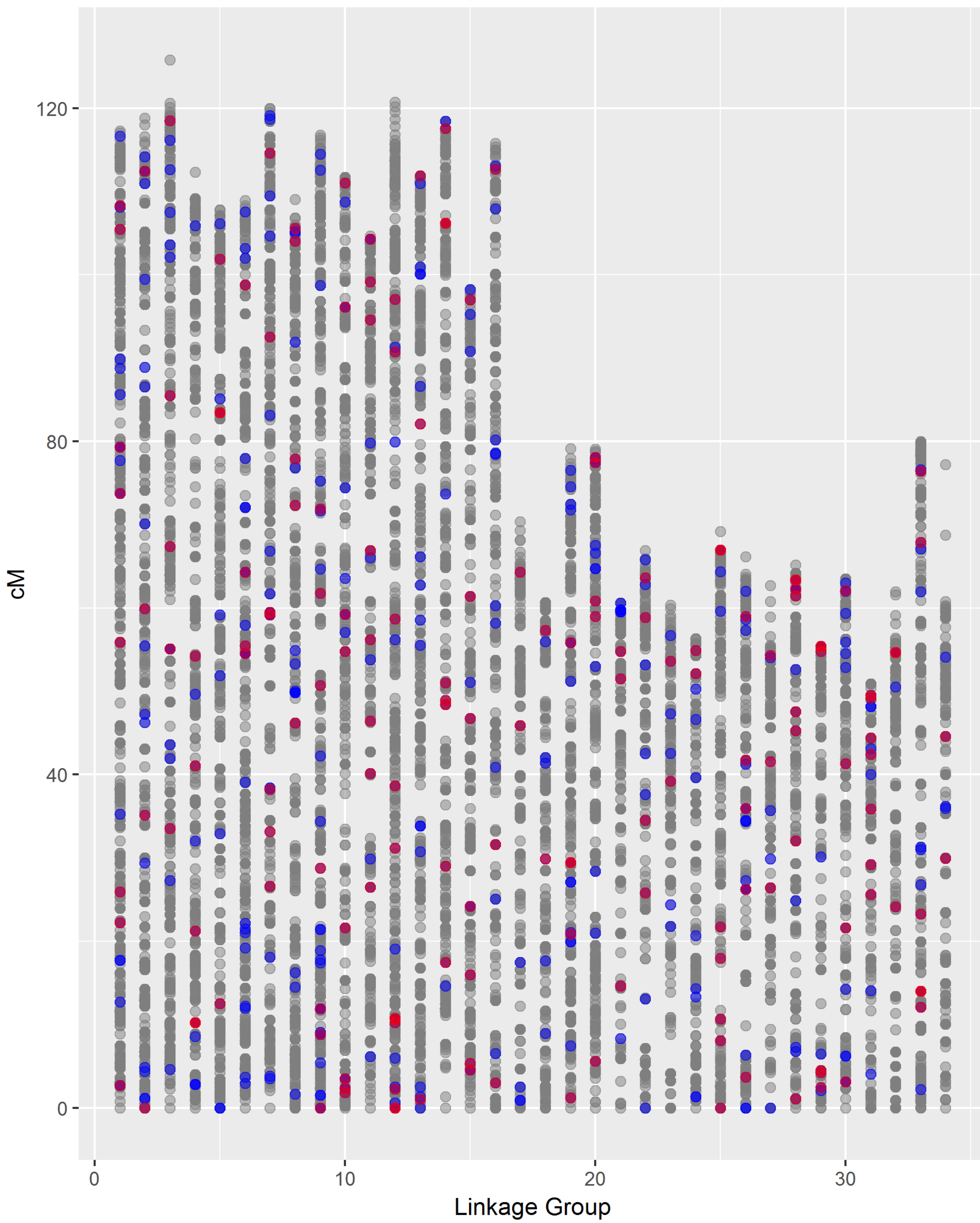

Figure 6.–Loci located on combined Washington-Alaska linkage map (grey), the 461 of 500 loci selected for panel development (blue) located on the linkage map, and the 171 incorporated in the final panel (red).

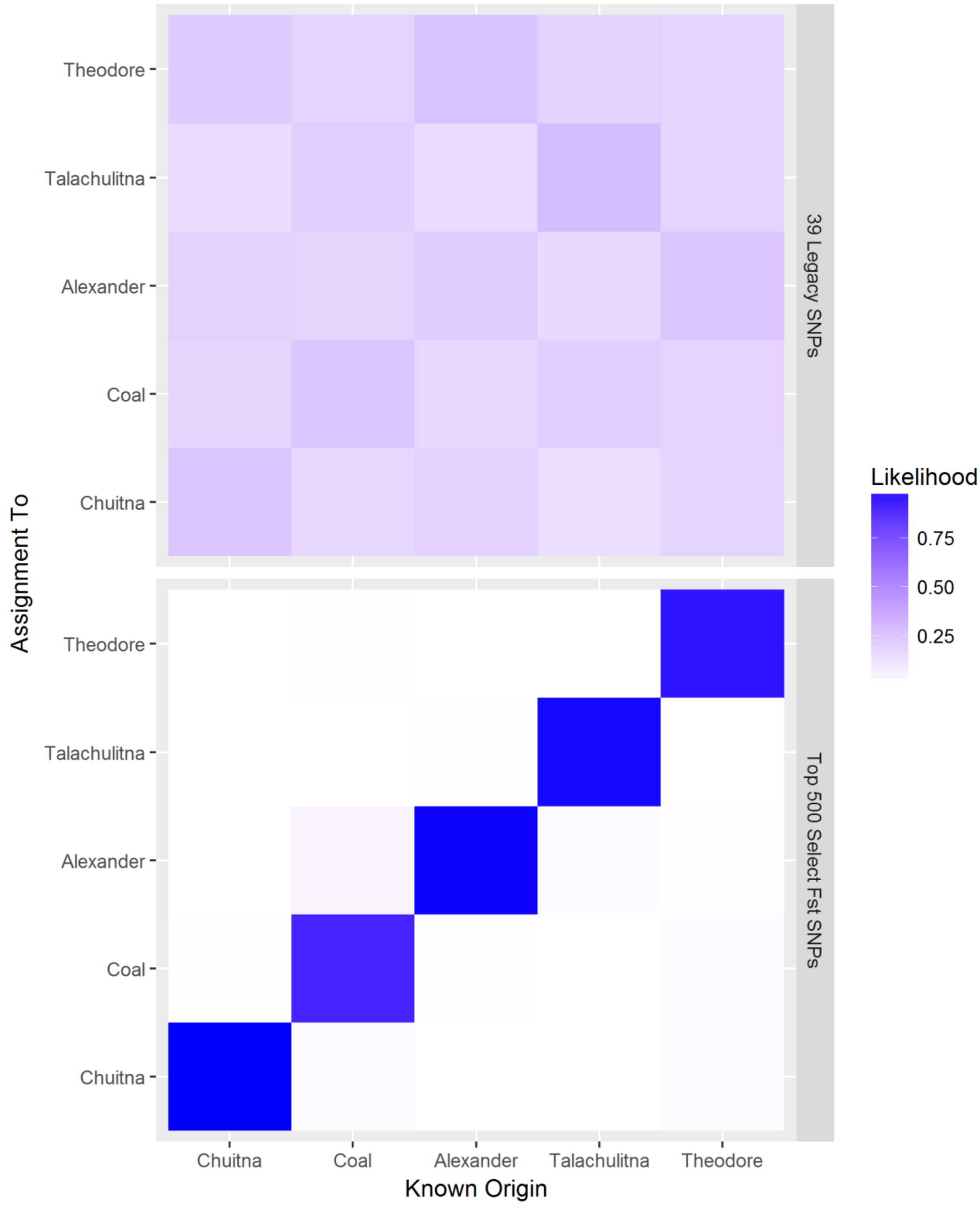

Figure 7.–Summary of mean genotype likelihood among all individuals from 5 populations of Cook Inlet Chinook salmon based upon the 39 legacy SNPs in the existing baseline (top) and 500 SNPs selected for panel development based upon F_{ST} (bottom). Probabilities off the diagonal indicate uncertainty in genetic assignment and provide indications of potential misallocation in mixed stock analysis.

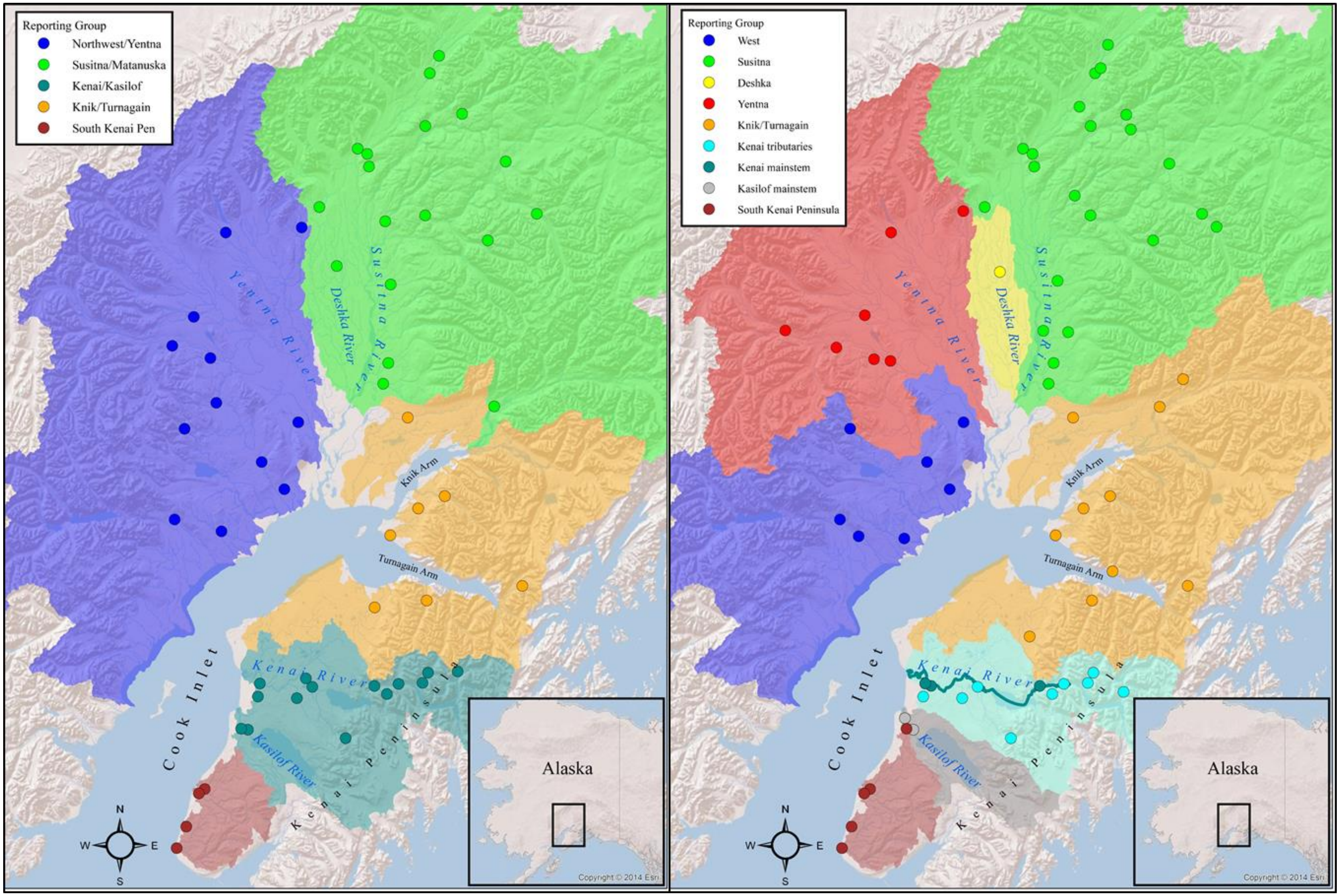

Figure 8.–Location and reporting group affiliation for 55 populations genotyped for the legacy 39 SNPs (left panel) and 66 populations genotyped for 414 SNPs (right panel) that include SNPs developed to better differentiate populations of Cook Inlet Chinook salmon (Appendix A).

APPENDIX A: SUMMARY OF 500 AMPLICONS SELECTED FOR PANEL DEVELOPMENT

Appendix A1.–Summary of the 500 amplicons selected for panel development including catalog ID, whether the amplicon was included in the final panel, purpose, linkage group, and position in centiMorgans (cM) for those amplicons located on the linkage map, and four F_{ST} values: among all populations (Overall), between Coal Creek and other populations (Coal), between Talachulitna River and other populations (Tala), and among the Chuitna and Theodore rivers and Alexander Creek (Cluster).

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
392	Yes	Cluster	Ots21	54.794	0.022	-0.004	0.000	0.048
418	Yes	Cluster	Ots26	59.01	0.035	0.012	-0.006	0.055
464	No	Coal	Ots19	19.935	0.026	0.071	-0.006	-0.005
839	No	Tala	Ots09	0	0.026	-0.001	0.062	0.002
904	No	Tala	Ots11	66.057	0.030	0.012	0.062	0.009
964	No	Tala	Ots01	85.593	0.024	-0.002	0.071	-0.007
1143	No	Tala	Ots02	47.24	0.028	0.000	0.069	0.003
1565	No	Cluster	Ots13	110.966	0.032	0.010	-0.005	0.061
1921	No	Cluster	NA	NA	0.019	-0.006	0.004	0.036
2844	No	Coal	Ots08	91.932	0.033	0.086	0.012	-0.003
2868	Yes	Coal	Ots13	111.909	0.034	0.080	0.038	-0.007
3283	No	Tala	Ots02	1.15	0.024	-0.002	0.074	-0.008
3358	No	Tala	Ots04	105.88	0.020	0.005	0.058	-0.005
3729	No	Coal	Ots09	11.891	0.026	0.078	0.002	-0.006
3830	Yes	Cluster	Ots26	41.737	0.034	0.031	-0.005	0.036
3884	Yes	Tala	Ots26	3.707	0.022	0.001	0.061	-0.001
4000	Yes	Cluster	Ots34	29.971	0.057	-0.005	0.054	0.056
4025	No	Cluster	Ots09	114.495	0.033	0.008	0.010	0.054
4299	No	Cluster	Ots16	78.595	0.021	-0.004	-0.007	0.048
4502	Yes	Cluster	Ots25	66.944	0.012	-0.007	-0.007	0.035
4783	No	Tala	Ots08	105.051	0.029	-0.002	0.061	-0.004
4809	No	Cluster	Ots15	90.78	0.027	0.000	0.023	0.036
5020	No	Cluster	Ots19	76.512	0.016	-0.006	-0.002	0.036
5164	No	Cluster	Ots19	19.935	0.023	-0.004	-0.005	0.054
5231	No	Tala	Ots09	8.825	0.032	-0.002	0.085	-0.004
5295	No	Coal	NA	NA	0.025	0.068	0.007	-0.002
5444	No	Cluster	Ots24	46.637	0.025	-0.006	0.010	0.049
5667	Yes	Coal	Ots21	51.482	0.023	0.071	-0.006	-0.010
5823	No	Tala	Ots11	6.139	0.022	-0.003	0.058	-0.003
5829	No	Cluster	Ots12	5.986	0.024	-0.003	-0.002	0.051
6287	No	Tala	Ots21	59.781	0.022	0.009	0.063	-0.008
6581	No	Cluster	Ots13	1.334	0.030	-0.006	0.019	0.044
6905	No	Coal	Ots22	53.192	0.046	0.065	-0.005	0.035
7195	No	Coal	Ots08	76.814	0.025	0.079	0.004	-0.008
7776	Yes	Tala	NA	NA	0.029	-0.006	0.072	-0.006
7932	No	Cluster	Ots05	32.909	0.024	0.002	0.004	0.041

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
8142	No	Cluster	NA	NA	0.015	-0.004	-0.005	0.037
8210	No	Tala	Ots13	30.748	0.047	-0.002	0.120	-0.004
8524	No	Cluster	Ots20	64.761	0.021	0.008	-0.007	0.039
8628	No	Coal	Ots13	2.49	0.027	0.074	-0.005	-0.006
8662	Yes	Tala	Ots12	0.001	0.047	-0.007	0.059	0.038
8834	Yes	Tala	Ots01	105.45	0.029	0.003	0.085	-0.007
8948	No	Cluster	Ots16	80.187	0.014	-0.003	-0.006	0.034
9374	No	Cluster	Ots26	26.226	0.033	-0.003	0.006	0.064
9427	No	Cluster	Ots31	14.054	0.024	-0.006	0.017	0.039
9688	Yes	Tala	Ots09	11.892	0.049	-0.007	0.096	0.019
9970	Yes	Cluster	Ots25	8.08	0.025	0.015	0.000	0.041
10049	Yes	Cluster	Ots12	90.77	0.025	0.004	-0.002	0.046
10092	No	Tala	Ots04	2.827	0.033	-0.006	0.060	0.013
10112	No	Cluster	Ots13	0	0.032	0.021	-0.006	0.040
10481	Yes	Coal	Ots32	24.139	0.058	0.072	0.025	0.007
10678	No	Coal	Ots08	1.637	0.041	0.065	0.004	0.020
10724	No	Coal	Ots13	33.841	0.036	0.098	0.000	-0.005
11531	No	Cluster	Ots19	71.756	0.013	-0.005	-0.005	0.036
11735	No	Coal	Ots10	59.209	0.029	0.074	-0.002	0.005
11975	No	Cluster	Ots08	49.84	0.035	-0.002	0.000	0.072
12417	No	Cluster	Ots09	5.433	0.032	-0.006	0.012	0.057
12524	Yes	Coal	Ots14	106.182	0.021	0.059	-0.005	-0.006
12688	Yes	Tala	NA	NA	0.019	-0.004	0.063	-0.010
12985	No	Cluster	Ots25	66.944	0.020	0.013	-0.003	0.035
13045	Yes	Cluster	Ots29	4.488	0.024	-0.004	0.006	0.043
13171	Yes	Tala	NA	NA	0.020	-0.007	0.064	-0.010
13409	No	Tala	Ots16	6.516	0.038	-0.007	0.078	0.011
13435	Yes	Coal	Ots30	62.05	0.034	0.095	0.000	-0.009
13711	Yes	Cluster	Ots11	26.473	0.017	0.005	-0.006	0.034
13755	Yes	Cluster	Ots15	5.324	0.036	0.035	-0.003	0.036
13797	No	Tala	Ots02	88.916	0.031	-0.004	0.078	-0.008
13938	No	Cluster	Ots16	25.042	0.040	0.020	-0.005	0.062
13946	No	Cluster	Ots21	59.515	0.017	-0.002	-0.006	0.039
14092	Yes	Coal	Ots12	31.148	0.027	0.060	0.000	0.007
14123	No	Cluster	Ots15	51.058	0.037	0.047	-0.006	0.042
14470	No	Cluster	Ots07	3.488	0.017	0.008	-0.004	0.035
14490	No	Cluster	Ots30	62.006	0.027	-0.006	-0.004	0.060
14778	No	Cluster	Ots06	72.068	0.042	0.030	0.008	0.035
15274	No	Coal	Ots21	8.287	0.027	0.068	-0.002	0.001

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
15287	Yes	Cluster	Ots28	45.293	0.017	-0.003	-0.006	0.038
15416	Yes	Tala	Ots04	10.214	0.029	0.002	0.060	-0.009
15859	Yes	Cluster	Ots14	29.009	0.026	0.000	0.005	0.050
15938	No	Cluster	Ots16	60.299	0.021	0.006	-0.006	0.041
16002	No	Cluster	Ots08	49.993	0.021	0.001	-0.001	0.041
16246	No	Cluster	Ots33	26.79	0.026	-0.004	0.019	0.036
16441	Yes	Cluster	Ots01	25.911	0.018	0.004	-0.007	0.037
16481	No	Cluster	Ots22	37.633	0.017	-0.005	0.006	0.034
16502	Yes	Cluster	Ots15	15.992	0.019	-0.006	-0.005	0.046
16523	Yes	Cluster	Ots16	31.603	0.044	0.012	-0.005	0.067
16625	Yes	Tala	Ots31	42.452	0.038	-0.003	0.073	0.003
16835	No	Cluster	Ots27	0	0.022	-0.005	-0.006	0.051
17027	Yes	Cluster	Ots03	67.385	0.029	-0.004	0.019	0.043
17420	Yes	Tala	Ots11	66.923	0.025	-0.005	0.065	-0.009
17472	No	Cluster	Ots02	114.172	0.028	0.031	-0.001	0.038
18078	No	Tala	Ots03	102.107	0.030	-0.002	0.062	-0.003
18162	No	Coal	Ots26	41.25	0.044	0.059	0.030	-0.009
18602	Yes	Tala	Ots23	39.24	0.021	0.004	0.068	-0.010
18808	No	Cluster	NA	NA	0.023	-0.009	0.007	0.050
18840	No	Coal	Ots21	59.515	0.034	0.074	-0.006	0.015
18987	No	Cluster	Ots10	96.146	0.033	-0.006	0.001	0.069
19139	Yes	Cluster	NA	NA	0.025	0.005	0.014	0.035
19423	Yes	Coal	Ots19	55.852	0.028	0.059	0.001	-0.006
19425	No	Cluster	Ots05	0	0.029	-0.006	-0.003	0.066
19707	Yes	Cluster	Ots10	59.218	0.023	0.020	-0.003	0.034
19835	No	Cluster	Ots23	24.384	0.015	0.000	-0.006	0.035
19849	No	Cluster	Ots24	1.304	0.031	-0.007	0.012	0.053
20110	Yes	Tala	Ots28	32.026	0.036	0.000	0.064	0.005
20153	No	Cluster	NA	NA	0.030	-0.005	-0.007	0.068
20168	No	Coal	Ots02	29.366	0.040	0.120	0.006	-0.010
20292	Yes	Tala	Ots28	61.449	0.022	-0.006	0.060	-0.006
20343	Yes	Cluster	Ots31	49.542	0.031	0.028	-0.005	0.034
20459	Yes	Coal	Ots33	12.092	0.023	0.059	-0.005	-0.006
20487	Yes	Tala	Ots09	8.825	0.036	0.005	0.083	0.006
20587	Yes	Cluster	Ots27	26.4	0.021	0.007	0.009	0.035
20670	No	Tala	Ots19	27.117	0.037	-0.007	0.090	-0.008
21105	No	Coal	Ots16	58.187	0.060	0.125	0.001	-0.008
21331	No	Cluster	Ots19	29.446	0.028	-0.004	0.006	0.048
21392	Yes	Coal	Ots33	13.985	0.035	0.073	0.000	-0.004

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
21678	Yes	Cluster	Ots08	46.162	0.018	-0.006	-0.003	0.039
21914	No	Cluster	Ots01	17.73	0.036	-0.006	0.001	0.073
22283	Yes	Tala	Ots22	34.529	0.021	0.000	0.062	-0.005
22426	Yes	Cluster	Ots11	94.591	0.023	-0.001	-0.003	0.042
23359	No	Cluster	Ots31	40	0.044	-0.006	0.034	0.066
23562	No	Cluster	Ots19	72.439	0.018	-0.004	-0.005	0.045
23565	Yes	Coal	Ots19	1.223	0.021	0.064	0.003	-0.006
23604	Yes	Cluster	Ots01	73.777	0.020	-0.007	0.000	0.041
23793	Yes	Cluster	Ots01	22.203	0.035	-0.002	0.031	0.046
23886	No	Cluster	Ots20	77.483	0.016	0.001	0.000	0.034
23913	No	Cluster	Ots10	57.066	0.023	0.003	0.014	0.036
24331	No	Cluster	Ots22	65.773	0.036	0.028	0.005	0.035
24458	Yes	Cluster	Ots27	41.555	0.025	0.007	-0.003	0.037
24567	No	Cluster	Ots24	50.295	0.013	-0.006	-0.005	0.036
24572	No	Cluster	Ots18	42.053	0.029	-0.006	0.018	0.041
24763	No	Coal	Ots26	0	0.020	0.069	-0.005	-0.011
24816	No	Tala	Ots09	18.89	0.032	-0.003	0.063	0.003
25055	Yes	Cluster	Ots34	44.565	0.021	-0.005	0.000	0.044
25234	Yes	Tala	Ots30	3.108	0.021	-0.002	0.060	-0.004
25273	Yes	Tala	Ots22	58.889	0.032	-0.005	0.083	-0.005
25625	No	Cluster	Ots06	57.928	0.014	-0.006	-0.004	0.037
25795	No	Coal	NA	NA	0.028	0.073	-0.006	-0.006
25876	Yes	Cluster	Ots22	25.828	0.046	-0.006	0.037	0.056
26045	No	Cluster	Ots30	63.011	0.043	0.017	0.016	0.039
26189	Yes	Coal	Ots30	41.307	0.031	0.085	0.007	-0.005
26263	No	Tala	Ots29	55.397	0.045	0.006	0.119	-0.009
26346	No	Coal	Ots06	54.56	0.032	0.077	0.034	-0.007
26449	No	Cluster	Ots05	51.858	0.022	0.000	-0.006	0.044
26506	No	Coal	Ots34	36.195	0.026	0.061	0.026	-0.004
26644	Yes	Cluster	Ots12	10.201	0.027	0.015	0.009	0.038
26657	Yes	Tala	Ots29	55.397	0.024	-0.005	0.068	-0.009
26757	Yes	Cluster	NA	NA	0.021	0.001	-0.003	0.043
27324	Yes	Coal	Ots15	46.736	0.029	0.061	0.029	0.001
28058	No	Coal	Ots10	63.539	0.030	0.071	-0.002	0.008
28238	Yes	Cluster	Ots28	63.368	0.017	0.001	-0.005	0.039
28526	No	Tala	NA	NA	0.030	-0.006	0.077	-0.009
28544	Yes	Tala	Ots04	10.214	0.028	-0.007	0.066	0.006
29090	No	Cluster	Ots09	98.732	0.039	0.037	0.027	0.038
29121	Yes	Cluster	Ots12	10.736	0.027	0.003	-0.006	0.055

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
29150	No	Coal	Ots28	62.268	0.021	0.058	-0.006	-0.004
29769	Yes	Tala	Ots08	104.05	0.031	-0.002	0.095	-0.011
30047	Yes	Cluster	Ots07	59.18	0.021	-0.006	0.011	0.036
30060	No	Coal	Ots07	18.07	0.032	0.064	0.001	-0.002
30120	No	Cluster	Ots29	2.05	0.016	-0.006	0.001	0.035
30208	No	Cluster	Ots26	6.346	0.022	-0.006	0.003	0.044
30345	Yes	Coal	Ots03	85.479	0.042	0.067	0.008	0.005
30562	Yes	Cluster	Ots25	66.944	0.044	0.027	-0.006	0.073
30759	Yes	Cluster	Ots07	38.19	0.025	-0.005	-0.004	0.057
30878	No	Cluster	Ots08	54.934	0.024	-0.005	0.017	0.037
31534	No	Coal	Ots29	6.456	0.030	0.066	-0.008	0.007
31577	Yes	Cluster	Ots03	33.553	0.026	0.030	-0.002	0.035
31599	No	Cluster	Ots08	16.243	0.014	-0.006	0.001	0.034
31796	Yes	Tala	Ots09	61.772	0.034	-0.006	0.068	0.013
31798	No	Tala	NA	NA	0.026	0.003	0.076	-0.009
32074	Yes	Cluster	Ots02	0	0.023	0.008	0.005	0.038
32126	No	Cluster	Ots15	4.532	0.023	-0.006	-0.005	0.050
32279	Yes	Cluster	Ots10	1.833	0.047	0.010	0.005	0.085
32287	Yes	Tala	Ots11	40.153	0.073	0.072	0.154	-0.010
32787	No	Tala	Ots04	8.571	0.036	0.027	0.090	-0.009
32804	No	Cluster	NA	NA	0.030	-0.003	0.041	0.036
32867	No	Cluster	Ots19	27.117	0.017	-0.009	-0.003	0.037
32984	No	Tala	Ots25	64.347	0.037	-0.003	0.058	0.015
33013	Yes	Cluster	Ots29	2.449	0.025	0.004	-0.006	0.054
33474	No	Cluster	NA	NA	0.021	0.013	-0.005	0.037
33502	No	Cluster	Ots14	73.721	0.045	0.045	0.021	0.049
33876	Yes	Tala	Ots21	14.627	0.034	0.018	0.074	0.003
34102	No	Cluster	Ots26	27.273	0.019	-0.001	0.001	0.039
34281	No	Cluster	Ots33	61.935	0.028	0.002	0.011	0.037
34486	No	Cluster	Ots06	19.163	0.029	0.014	-0.006	0.052
34631	No	Cluster	Ots07	59.18	0.015	-0.001	0.000	0.035
34797	No	Cluster	Ots28	63.398	0.018	0.003	-0.002	0.035
34802	Yes	Cluster	Ots27	54.3	0.030	0.029	0.009	0.034
35239	Yes	Cluster	Ots24	52.111	0.028	0.012	-0.003	0.051
35479	No	Coal	Ots20	52.973	0.020	0.064	-0.002	-0.008
35625	No	Cluster	Ots22	42.556	0.020	-0.002	0.002	0.040
35949	Yes	Cluster	Ots18	57.342	0.038	0.011	0.011	0.049
35997	No	Cluster	Ots02	46.284	0.038	0.026	-0.001	0.041
36007	No	Cluster	Ots26	34.513	0.018	-0.007	-0.006	0.048

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
36202	Yes	Cluster	Ots25	17.979	0.020	-0.006	-0.004	0.045
36203	No	Cluster	Ots13	100.914	0.019	-0.006	0.007	0.036
36740	No	Cluster	Ots03	112.652	0.013	0.003	-0.007	0.036
36801	No	Cluster	Ots01	116.607	0.022	0.002	0.009	0.037
36833	No	Cluster	Ots30	14.268	0.023	-0.002	0.016	0.037
36916	Yes	Cluster	Ots10	21.59	0.043	0.006	0.013	0.073
36932	No	Cluster	Ots28	24.845	0.049	0.003	0.036	0.057
37035	Yes	Tala	Ots12	97.033	0.034	0.010	0.087	-0.004
37275	Yes	Cluster	Ots07	59.46	0.032	-0.002	0.006	0.059
37661	Yes	Cluster	Ots10	54.769	0.018	-0.006	-0.001	0.037
37714	No	Cluster	Ots12	0.001	0.083	-0.007	0.030	0.118
37744	Yes	Tala	Ots02	59.883	0.021	-0.004	0.065	-0.008
38074	No	Cluster	Ots29	55.397	0.017	-0.004	-0.005	0.037
38104	Yes	Tala	Ots05	83.466	0.042	0.016	0.104	-0.006
38331	Yes	Cluster	Ots28	47.571	0.027	-0.006	-0.005	0.061
38337	Yes	Cluster	Ots09	71.92	0.016	0.001	-0.006	0.035
38515	No	Cluster	Ots02	55.469	0.020	0.006	-0.005	0.039
38531	No	Tala	Ots15	98.243	0.038	-0.003	0.063	0.023
38689	No	Cluster	Ots06	21.477	0.021	-0.005	0.007	0.036
38946	No	Tala	Ots13	33.841	0.045	0.001	0.103	0.008
39314	No	Cluster	Ots05	59.157	0.014	-0.005	-0.004	0.037
39584	No	Cluster	Ots06	103.168	0.014	-0.005	-0.007	0.038
39725	No	Tala	Ots03	43.62	0.036	0.008	0.082	0.006
40026	No	Cluster	Ots31	48.19	0.034	0.020	0.015	0.047
40086	Yes	Tala	Ots20	77.483	0.049	0.010	0.083	0.004
40163	Yes	Cluster	Ots09	50.704	0.025	0.005	-0.008	0.051
40389	No	Cluster	Ots34	35.937	0.030	-0.005	-0.005	0.060
40477	Yes	Cluster	Ots19	29.446	0.027	-0.003	0.005	0.049
40562	No	Tala	NA	NA	0.021	-0.007	0.059	-0.005
40588	Yes	Cluster	Ots01	79.29	0.030	-0.006	-0.003	0.062
40978	No	Tala	Ots04	2.827	0.034	-0.001	0.064	0.009
41051	No	Cluster	Ots03	107.494	0.025	-0.004	0.012	0.035
41272	No	Cluster	Ots20	64.761	0.018	0.000	-0.003	0.034
41732	No	Coal	Ots26	62.001	0.039	0.102	0.017	-0.006
42036	No	Cluster	Ots08	105.198	0.026	0.015	-0.006	0.043
42060	No	Cluster	Ots02	86.529	0.028	0.002	0.017	0.038
42112	No	Cluster	Ots16	78.437	0.015	-0.006	-0.003	0.036
42562	Yes	Cluster	Ots20	78.088	0.017	-0.003	-0.005	0.042
42766	No	Cluster	NA	NA	0.025	0.002	-0.004	0.050

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
42851	Yes	Cluster	Ots01	2.664	0.034	-0.004	-0.003	0.076
42863	No	Cluster	Ots06	39.069	0.027	0.001	0.004	0.040
42864	Yes	Coal	Ots05	12.539	0.020	0.060	-0.003	-0.003
43082	Yes	Tala	Ots32	54.648	0.047	0.008	0.113	0.001
43086	Yes	Cluster	NA	NA	0.015	-0.006	-0.006	0.038
43187	No	Cluster	Ots03	103.607	0.014	-0.006	-0.004	0.036
43322	No	Tala	Ots09	21.39	0.061	-0.007	0.124	0.009
43828	No	Cluster	Ots13	55.501	0.027	0.010	-0.005	0.054
44834	Yes	Coal	Ots33	23.254	0.040	0.098	-0.006	-0.003
44889	Yes	Coal	Ots02	112.405	0.027	0.075	0.009	0.001
45063	Yes	Cluster	Ots13	82.075	0.030	-0.006	0.020	0.040
45344	No	Cluster	Ots02	110.927	0.015	-0.002	-0.005	0.037
45623	No	Cluster	Ots06	54.56	0.026	-0.004		0.044
46175	No	Tala	Ots19	7.444	0.029	-0.007	0.073	0.000
46327	No	Cluster	Ots18	55.966	0.022	0.005	-0.007	0.035
46345	No	Tala	Ots09	17.413	0.046	0.024	0.082	0.022
46755	No	Cluster	Ots12	56.195	0.013	-0.007	-0.006	0.037
46842	Yes	Cluster	Ots31	49.248	0.023	0.000	-0.003	0.042
46901	No	Coal	Ots11	53.786	0.044	0.105	0.006	0.009
46922	No	Tala	Ots21	59.781	0.041	0.006	0.073	-0.009
47191	Yes	Coal	Ots17	45.911	0.077	0.117	0.020	0.010
47412	No	Cluster	Ots12	10.743	0.015	0.004	-0.006	0.034
47758	No	Cluster	Ots06	21.021	0.026	0.000	0.003	0.046
48032	Yes	Cluster	Ots10	96.146	0.024	-0.003	-0.001	0.052
48148	No	Cluster	NA	NA	0.020	0.003	-0.007	0.051
48167	No	Cluster	Ots09	42.244	0.018	-0.005	-0.005	0.041
48209	No	Cluster	Ots08	53.325	0.027	0.002	0.002	0.037
48478	No	Cluster	Ots02	4.341	0.024	0.004	-0.004	0.044
48610	No	Cluster	Ots16	107.897	0.011	-0.008	-0.007	0.034
48649	Yes	Coal	Ots16	3.019	0.037	0.075	0.036	0.001
48852	No	Cluster	NA	NA	0.018	0.005	-0.001	0.038
48855	Yes	Coal	Ots12	38.635	0.026	0.062	-0.003	-0.004
49015	No	Cluster	Ots08	49.84	0.014	-0.006	-0.006	0.039
49326	No	Coal	Ots31	43.095	0.020	0.058	-0.007	-0.005
49553	No	Cluster	Ots09	64.678	0.019	-0.004	0.008	0.037
49621	No	Cluster	Ots30	59.329	0.033	0.012	0.014	0.049
50114	No	Cluster	Ots15	95.28	0.022	0.016	-0.004	0.036
50399	No	Tala	NA	NA	0.042	0.017	0.062	-0.005
50458	Yes	Cluster	Ots06	98.791	0.030	0.016	-0.001	0.046

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
50525	No	Tala	Ots30	55.88	0.063	0.072	0.077	0.028
50811	No	Cluster	Ots19	74.542	0.019	0.005	-0.006	0.038
50837	No	Coal	Ots10	3.487	0.024	0.066	-0.006	-0.005
51032	Yes	Tala	Ots14	48.907	0.068	0.069	0.122	0.005
51278	No	Cluster	Ots05	106.123	0.026	0.008	0.003	0.047
51364	No	Cluster	Ots06	101.99	0.031	0.010	0.007	0.042
51701	No	Tala	Ots26	35.941	0.035	0.009	0.058	-0.004
51844	No	Cluster	NA	NA	0.018	-0.007	-0.005	0.043
52242	Yes	Coal	NA	NA	0.030	0.078	-0.004	0.003
52493	No	Cluster	Ots10	74.451	0.015	-0.005	-0.005	0.039
52671	No	Cluster	NA	NA	0.019	-0.007	-0.006	0.048
53049	No	Tala	Ots06	77.957	0.020	-0.003	0.061	-0.008
53050	Yes	Tala	Ots12	2.25	0.047	0.004	0.123	-0.009
53102	No	Cluster	Ots12	19.08	0.013	-0.006	-0.007	0.036
53513	Yes	Cluster	Ots15	4.532	0.031	-0.001	-0.005	0.068
53612	No	Cluster	Ots01	79.284	0.014	-0.002	-0.004	0.039
54163	No	Cluster	Ots33	31.272	0.032	-0.004	0.032	0.042
54208	No	Coal	Ots06	72.068	0.020	0.067	-0.004	-0.009
54236	No	Cluster	NA	NA	0.053	-0.006	0.000	0.108
54388	No	Cluster	Ots17	2.505	0.033	-0.006	0.023	0.042
54614	Yes	Coal	Ots11	99.152	0.025	0.061	-0.001	0.007
54653	Yes	Cluster	Ots07	33.166	0.044	-0.006	0.014	0.073
54765	No	Cluster	Ots04	32.055	0.032	-0.005	0.041	0.037
55425	Yes	Cluster	Ots04	54.241	0.028	0.003	-0.004	0.047
55538	Yes	Coal	NA	NA	0.025	0.064	-0.005	-0.006
55571	Yes	Tala	Ots11	46.438	0.035	-0.006	0.077	0.010
55829	No	Cluster	Ots33	76.596	0.017	-0.001	-0.001	0.037
56369	No	Cluster	Ots03	27.308	0.024	-0.006	0.002	0.043
56440	No	Cluster	Ots19	21.024	0.030	0.036	-0.002	0.035
56668	No	Tala	Ots03	55.092	0.031	0.002	0.084	-0.006
56955	No	Coal	Ots01	35.283	0.026	0.060	0.028	-0.004
57006	Yes	Coal	Ots05	101.891	0.059	0.138	-0.004	-0.008
57041	No	Cluster	Ots18	41.355	0.039	0.011	0.004	0.048
57266	No	Cluster	Ots03	41.924	0.014	-0.005	-0.006	0.038
57475	No	Cluster	Ots08	14.477	0.014	-0.005	-0.006	0.035
57521	No	Cluster	Ots21	60.579	0.035	-0.002	0.006	0.056
57654	Yes	Coal	Ots28	63.398	0.057	0.093	0.050	0.022
57669	Yes	Cluster	Ots19	29.446	0.035	0.001	0.007	0.059
57815	No	Cluster	Ots09	112.557	0.024	0.016	0.000	0.035

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
57941	No	Tala	Ots11	29.878	0.030	-0.004	0.066	-0.004
58089	No	Coal	Ots06	2.919	0.030	0.066	-0.005	0.005
58404	No	Coal	Ots33	2.201	0.030	0.065	-0.006	0.008
59380	No	Coal	Ots06	22.137	0.023	0.060	0.005	-0.002
59572	Yes	Coal	Ots30	21.618	0.034	0.078	-0.005	0.010
59601	No	Cluster	Ots01	73.777	0.010	-0.001	-0.007	0.034
59667	Yes	Tala	Ots31	44.414	0.026	-0.006	0.070	-0.007
59888	Yes	Tala	Ots09	0	0.025	-0.006	0.067	-0.010
60118	No	Cluster	Ots31	48.19	0.029	0.007	0.016	0.043
60124	Yes	Cluster	Ots03	55.092	0.046	0.001	0.026	0.067
60132	Yes	Coal	Ots10	2.41	0.025	0.070	-0.005	-0.007
60283	No	Cluster	Ots26	58.615	0.029	-0.002	-0.009	0.062
60285	Yes	Cluster	Ots14	17.484	0.020	-0.002	-0.005	0.040
60332	Yes	Tala	Ots12	0.001	0.034	0.004	0.094	-0.008
60347	No	Tala	Ots33	31.021	0.053	0.004	0.095	-0.001
60582	No	Cluster	Ots16	40.867	0.018	0.004	-0.008	0.035
61154	No	Cluster	NA	NA	0.028	0.002	-0.006	0.056
61296	No	Cluster	Ots18	17.672	0.038	0.016	0.020	0.048
61345	Yes	Cluster	Ots18	29.883	0.032	-0.006	0.010	0.049
61488	No	Cluster	Ots23	21.761	0.023	0.004	0.008	0.040
61810	No	Cluster	Ots09	11.891	0.024	-0.007	-0.007	0.057
61893	No	Cluster	Ots09	34.385	0.039	0.035	0.024	0.041
62017	Yes	Tala	NA	NA	0.027	-0.006	0.058	0.008
62284	No	Tala	Ots31	4.064	0.033	-0.003	0.074	-0.008
62367	No	Cluster	Ots27	29.909	0.015	0.001	-0.006	0.036
62758	No	Cluster	Ots30	54.505	0.030	0.003	-0.006	0.057
63065	Yes	Cluster	Ots29	4.209	0.029	0.009	-0.006	0.056
63105	Yes	Tala	Ots09	28.792	0.037	0.006	0.066	0.001
63254	No	Tala	NA	NA	0.022	-0.007	0.067	-0.009
63269	No	Cluster	Ots19	51.2	0.017	-0.001	-0.004	0.038
64082	Yes	Coal	Ots33	76.386	0.023	0.072	-0.001	-0.009
64288	Yes	Coal	Ots29	55.397	0.024	0.060	-0.005	0.002
64291	Yes	Coal	NA	NA	0.025	0.071	0.001	-0.002
64319	No	Tala	Ots26	26.226	0.037	-0.004	0.082	0.010
64661	No	Cluster	Ots20	28.391	0.022	-0.004	-0.002	0.051
64738	No	Cluster	Ots26	57.318	0.044	-0.006	-0.001	0.087
64886	No	Cluster	Ots20	66.564	0.014	-0.005	-0.005	0.036
65000	Yes	Coal	Ots10	3.487	0.033	0.087	-0.006	-0.005
65113	No	Cluster	Ots24	20.675	0.014	-0.006	-0.005	0.035

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
65234	No	Cluster	Ots06	64.324	0.032	0.014	-0.005	0.048
65356	No	Cluster	Ots05	85.096	0.020	0.000	-0.006	0.045
65466	Yes	Cluster	NA	NA	0.035	0.019	-0.004	0.043
65625	No	Cluster	Ots08	49.84	0.020	-0.006	-0.004	0.051
65879	No	Cluster	Ots01	17.728	0.016	0.001	-0.006	0.035
66125	No	Cluster	Ots07	59.18	0.019	-0.004	0.001	0.040
66360	Yes	Cluster	Ots12	10.736	0.033	0.041	-0.005	0.040
66433	Yes	Coal	Ots28	62.268	0.026	0.063	-0.006	0.001
66508	No	Cluster	Ots01	2.664	0.013	-0.005	-0.006	0.034
66791	Yes	Coal	Ots25	10.678	0.030	0.064	0.001	-0.006
66848	Yes	Tala	NA	NA	0.026	-0.006	0.066	0.002
67011	No	Cluster	Ots12	2.478	0.022	0.014	-0.005	0.034
67272	No	Coal	Ots24	13.348	0.062	0.083	0.042	0.041
67562	No	Cluster	Ots13	100.079	0.026	0.006	-0.005	0.043
67940	No	Cluster	Ots30	6.194	0.019	-0.006	0.000	0.041
68038	No	Cluster	Ots03	4.581	0.040	0.038	0.003	0.034
68561	No	Cluster	Ots28	6.768	0.015	-0.007	0.001	0.037
68590	No	Tala	Ots01	88.776	0.030	0.002	0.084	-0.004
68831	Yes	Cluster	Ots24	54.929	0.018	0.000	0.004	0.034
69027	Yes	Tala	Ots14	48.435	0.077	0.100	0.105	0.015
69145	No	Tala	Ots07	61.711	0.021	-0.004	0.064	-0.008
69475	No	Coal	NA	NA	0.038	0.109	0.002	-0.009
69633	No	Cluster	Ots33	67.106	0.033	0.051	-0.005	0.035
70054	No	Cluster	Ots13	58.524	0.016	-0.004	0.001	0.036
70063	Yes	Cluster	Ots31	25.621	0.036	0.005	0.019	0.037
70093	No	Cluster	Ots26	34.408	0.013	-0.005	-0.005	0.035
70444	No	Cluster	Ots05	0	0.022	0.008	-0.006	0.037
70623	No	Cluster	Ots20	67.508	0.029	-0.006	0.019	0.042
70980	No	Cluster	Ots13	62.729	0.017	0.003	-0.005	0.036
71081	No	Tala	NA	NA	0.041	-0.004	0.109	-0.010
71381	No	Coal	Ots09	9.05	0.036	0.079	-0.009	0.008
71514	Yes	Coal	Ots26	35.941	0.033	0.079	-0.001	0.004
71757	No	Coal	Ots07	119.092	0.061	0.101	0.001	0.048
71891	No	Cluster	Ots11	79.788	0.028	0.026	-0.001	0.041
72338	No	Coal	Ots32	54.648	0.037	0.091	-0.006	-0.005
72607	No	Cluster	Ots30	52.847	0.015	0.000	-0.006	0.034
72690	No	Cluster	NA	NA	0.033	0.046	-0.006	0.034
72961	Yes	Tala	Ots29	54.796	0.028	-0.001	0.081	-0.007
73090	No	Cluster	Ots06	11.973	0.024	-0.006	0.006	0.045

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
73097	Yes	Cluster	Ots06	54.56	0.033	-0.003	0.024	0.045
73140	Yes	Tala	Ots25	0	0.039	-0.003	0.079	0.000
73167	No	Coal	Ots24	1.304	0.024	0.059	-0.005	-0.002
73366	Yes	Cluster	Ots07	114.598	0.027	0.008	0.001	0.046
73402	Yes	Coal	Ots04	21.228	0.025	0.058	-0.005	0.002
73578	No	Cluster	Ots23	47.29	0.024	0.009	-0.007	0.044
73593	No	Cluster	Ots07	109.462	0.017	-0.004	-0.003	0.041
73604	Yes	Coal	Ots15	61.404	0.027	0.073	0.018	-0.008
73786	Yes	Tala	Ots08	105.594	0.036	0.012	0.099	-0.008
74404	Yes	Cluster	Ots11	104.282	0.043	0.004	0.016	0.060
74511	Yes	Tala	Ots01	108.299	0.039	0.003	0.075	-0.003
74604	No	Cluster	Ots30	3.108	0.025	0.026	-0.004	0.039
74833	Yes	Coal	Ots33	13.985	0.035	0.077	-0.005	0.009
74972	No	Cluster	Ots01	12.723	0.043	-0.005	0.016	0.061
75069	Yes	Cluster	Ots08	77.862	0.034	0.008	0.014	0.053
75464	No	Cluster	Ots12	91.281	0.026	0.001	-0.002	0.047
75627	Yes	Coal	Ots20	5.604	0.045	0.070	0.019	-0.007
75885	Yes	Cluster	Ots31	35.869	0.037	0.002	-0.005	0.079
76197	Yes	Coal	Ots11	56.201	0.045	0.063	0.018	-0.003
76279	No	Coal	Ots06	12.189	0.037	0.076	0.046	-0.008
77086	No	Tala	Ots01	108.089	0.023	-0.006	0.066	-0.007
77831	Yes	Cluster	Ots19	20.875	0.037	0.043	0.002	0.043
78184	No	Coal	Ots27	35.689	0.029	0.072	-0.004	-0.010
78324	No	Cluster	Ots07	83.171	0.022	-0.005		0.035
79551	No	Cluster	NA	NA	0.017	-0.007	0.000	0.040
79860	No	Cluster	Ots07	38.386	0.026	-0.002	0.023	0.040
80431	Yes	Cluster	Ots16	112.704	0.024	-0.005	0.004	0.045
80510	Yes	Cluster	Ots14	117.584	0.027	-0.004	0.005	0.055
80789	No	Cluster	Ots18	8.902	0.032	-0.005	0.031	0.041
80839	No	Coal	Ots22	0	0.045	0.105	0.020	0.001
81084	Yes	Cluster	Ots12	58.649	0.019	-0.005	-0.004	0.038
81372	No	Cluster	Ots28	52.631	0.021	-0.007	-0.002	0.044
81543	Yes	Tala	Ots14	106.182	0.026	0.007	0.061	0.004
81691	No	Tala	Ots07	3.831	0.029	-0.007	0.075	0.000
81746	No	Tala	Ots12	0.568	0.032	-0.007	0.067	0.009
81807	No	Cluster	Ots20	20.962	0.027	-0.005	0.015	0.041
81927	Yes	Tala	Ots25	21.697	0.030	0.004	0.082	-0.004
82047	Yes	Coal	Ots07	26.583	0.031	0.060	-0.003	0.026
82247	Yes	Cluster	Ots06	64.32	0.021	-0.006	-0.006	0.050

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
82609	No	Cluster	Ots13	66.128	0.030	0.001	0.032	0.040
82889	Yes	Tala	Ots05	83.466	0.039	0.020	0.098	-0.008
83004	Yes	Coal	Ots08	72.304	0.035	0.083	-0.005	-0.002
83141	No	Cluster	Ots07	66.802	0.023	0.020	-0.005	0.038
83331	No	Coal	Ots19	55.852	0.023	0.071	-0.004	-0.007
83466	No	Coal	Ots06	3.702	0.027	0.058	-0.007	0.012
83732	Yes	Tala	Ots20	58.975	0.032	-0.003	0.060	0.016
83915	No	Cluster	Ots01	79.288	0.020	-0.007	0.002	0.045
84134	No	Tala	Ots02	4.828	0.021	-0.005	0.061	-0.004
84318	Yes	Cluster	Ots31	29.224	0.022	-0.006	-0.006	0.053
84598	Yes	Cluster	Ots23	53.596	0.017	-0.002	0.001	0.036
84873	No	Coal	Ots10	2.41	0.034	0.091	-0.007	-0.006
85155	No	Cluster	Ots04	2.827	0.022	0.000	-0.004	0.048
85978	No	Cluster	Ots24	14.28	0.016	-0.006	-0.006	0.040
86211	Yes	Cluster	Ots28	1.077	0.029	0.000	0.009	0.053
86395	No	Cluster	Ots07	59.46	0.017	-0.003	0.003	0.036
86912	No	Cluster	NA	NA	0.019	0.006	0.003	0.034
86916	No	Cluster	Ots13	86.624	0.014	-0.006	-0.006	0.037
87106	No	Coal	Ots29	30.096	0.058	0.067	0.044	-0.007
87297	No	Coal	Ots24	39.64	0.050	0.070	-0.004	0.031
87524	No	Cluster	Ots09	1.501	0.016	-0.005	-0.006	0.064
88246	No	Cluster	Ots23	42.56	0.036	0.010	-0.007	0.061
88339	No	Coal	Ots09	21.382	0.030	0.084	-0.001	-0.003
88890	No	Cluster	Ots07	104.665	0.026	-0.007	-0.004	0.058
88897	Yes	Cluster	Ots04	41.073	0.051	0.043	0.043	0.047
88914	No	Coal	Ots09	71.672	0.024	0.067	0.003	-0.002
89199	No	Cluster	Ots26	34.403	0.020	-0.003	0.004	0.040
89364	No	Tala	Ots13	100.079	0.044	-0.007	0.099	0.005
89377	No	Cluster	Ots02	70.088	0.018	-0.002	0.003	0.038
89430	No	Cluster	Ots28	62.267	0.030	0.018	-0.003	0.039
89492	No	Tala	NA	NA	0.025	0.004	0.065	0.000
89746	No	Cluster	Ots30	6.194	0.033	0.025	0.005	0.048
89884	No	Cluster	Ots17	17.473	0.018	-0.006	-0.006	0.044
91227	No	Cluster	Ots02	112.405	0.028	0.005	0.009	0.040
92126	No	Cluster	Ots10	108.726	0.044	0.025	0.050	0.037
92666	Yes	Coal	Ots15	24.224	0.023	0.071	-0.001	-0.007
92901	Yes	Cluster	Ots01	55.907	0.026	0.007	-0.003	0.038
93170	Yes	Coal	Ots13	1.049	0.026	0.075	-0.006	-0.004
93789	Yes	Cluster	Ots33	67.899	0.031	0.012	0.015	0.044

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
95280	No	Cluster	Ots02	1.145	0.017	0.004	-0.006	0.035
95549	No	Coal	Ots26	0	0.036	0.072	-0.009	0.017
98206	No	Cluster	Ots22	13.096	0.014	-0.009	-0.007	0.036
98255	Yes	Coal	Ots03	118.476	0.031	0.059	-0.003	0.008
98734	No	Coal	Ots01	77.714	0.037	0.077	0.033	0.003
99972	No	Cluster	Ots28	7.233	0.022	0.006	-0.006	0.042
100237	Yes	Cluster	Ots02	35.107	0.020	0.002	0.001	0.040
100529	No	Coal	NA	NA	0.041	0.066	-0.010	0.031
100808	No	Cluster	Ots12	79.915	0.020	0.001	-0.004	0.035
101818	Yes	Cluster	Ots15	96.97	0.070	0.015	0.024	0.081
103380	Yes	Coal	Ots10	111.027	0.031	0.069	-0.004	0.005
103394	Yes	Tala	Ots14	51.038	0.051	-0.005	0.131	0.001
103395	No	Cluster	Ots03	116.137	0.015	-0.002	-0.006	0.037
103720	No	Cluster	Ots09	17.761	0.022	0.004	0.011	0.035
104500	No	Cluster	Ots09	1.501	0.006	-0.007	-0.004	0.034
104966	No	Cluster	Ots09	75.257	0.021	-0.007	0.010	0.037
105150	Yes	Tala	Ots22	63.623	0.023	-0.004	0.064	-0.003
105314	No	Cluster	Ots22	62.822	0.028	0.006	0.008	0.038
106168	No	Cluster	Ots21	59.515	0.015	-0.002	-0.006	0.035
107496	No	Cluster	Ots25	59.619	0.022	-0.006	-0.006	0.054
107933	No	Cluster	Ots15	24.176	0.017	-0.006	-0.002	0.041
108943	Yes	Cluster	Ots06	55.45	0.025	0.001	-0.002	0.052
109081	No	Coal	Ots17	0.911	0.022	0.066	0.002	-0.003
109094	No	Cluster	Ots16	113.079	0.034	0.015	0.011	0.034
109411	Yes	Tala	Ots26	26.226	0.030	-0.005	0.067	0.007
109670	No	Cluster	Ots02	99.463	0.021	-0.005	0.014	0.036
110335	No	Cluster	Ots20	77.966	0.023	-0.005	-0.004	0.055
111118	No	Cluster	Ots17	0.914	0.032	-0.001	0.008	0.059
111430	Yes	Cluster	Ots20	60.818	0.053	0.012	0.011	0.071
111926	No	Coal	Ots01	89.893	0.037	0.065	-0.006	0.024

-continued-

Catalog ID	Included in final panel	Purpose	LG	cM	Overall	Coal	Tala	Cluster
112054	No	Coal	Ots04	49.656	0.031	0.066	0.011	0.008
112461	Yes	Tala	Ots07	92.526	0.040	-0.007	0.062	0.021
113122	No	Cluster	Ots11	104.282	0.041	0.010	0.014	0.043
113376	No	Cluster	Ots14	118.431	0.020	-0.002	0.000	0.041
118266	No	Coal	Ots07	118.692	0.082	0.158	-0.001	0.006
Total Yes	183							
Total on linkage map			461					
Overall average					0.029	0.018	0.018	0.027
Total Cluster		300						
Cluster average					0.026	0.003	0.003	0.045
Total Coal		100						
Coal average					0.034	0.075	0.004	0.001
Total Tala		100						
Tala average					0.034	0.003	0.077	0.000