Development of a standardized suite of microsatellite loci to be used in the establishment of a chum salmon baseline for southern British Columbia and Washington

Denise K. Hawkins, Sewall F. Young

Washington Department of Fish and Wildlife
600 Capitol Way North
Olympia, Washington 98501

Terry D. Beacham, Mike H. Wetklo

Department of Fisheries and Oceans
Pacific Biological Station
Nanaimo, BC
Canada V9T 6N7

USA

June 2005

A project funded by the Southern Boundary Restoration and Enhancement Fund 2004-2005.

Contents

List of Tables

List of Figures

Abstract

The objective of the project was to reach agreement on the microsatellite loci to include in development of a standardized baseline for chum salmon, and to demonstrate that concordance between laboratories is possible in identification of alleles observed at each locus. Staff in genetics laboratories from the Canadian Department of Fisheries and Oceans and the Washington Department of Fish and Wildlife reached agreement on 15 microsatellite loci to include in the development of a standardized baseline. They also demonstrated that concordance in allele identification between laboratories was greater than 99.5% for all loci in both a developmental survey of genetic variation and in a blind sample test. It is now possible to develop a standardized microsatellite baseline for stock identification applications in chum salmon fisheries in southern British Columbia and Washington.

Introduction

An ongoing data need for managing chum salmon fisheries and harvests throughout the Southern Boundary area is estimation of stock contributions (and exploitation rates) for major mixed-stock fishery harvests in southern B.C. and Washington. While allozyme (protein) analysis can provide informative estimates of contributions for genetically divergent groups of populations it does not have the resolving power necessary for stock-specific analyses (Beacham et al. 1987). Based on preliminary studies of chum and other species of Pacific salmon, microsatellite DNA shows considerable promise for increasing the resolution of such analyses to the level of individual stocks (Beacham et al. 2005). However, the absence of a standardized approach for DNA analysis by DFO and WDFW (the two agencies with primary fishery management responsibility in this region) has prevented the development of a shared DNA baseline for addressing PST-level issues between Canada and the U.S. This report outlines the progress in developing a shared baseline for chum salmon: the establishment of a suite of 15 standardized microsatellite loci to be used in generating baseline data. This baseline will be shared between Washington and British Columbia fishery management agencies for application to estimation of stock composition in mixed-stock fishery samples.

Methods and Materials

DNA extraction by DFO and WDFW was conducted using either a commercially available 96-well format extraction kit or another protocol such as chelex extraction. Selected microsatellite DNA loci of interest were amplified via the polymerase chain reaction (PCR) using fluorescently labeled primers. The DFO laboratory analyzed all amplified fragments using an ABI-377 slab gel automated DNA sequencer in a 96-well format. The WDFW laboratory analyzed all amplified fragments using an ABI 3730 capillary automated DNA Analyzer in a 48-capillary format. Analytical software for each instrument was used to collect and analyze the raw data to estimate the sizes (in base pairs) of the fragments at each locus. Estimated allele sizes were then compared among laboratories in an initial standardization process. Allele names were derived from the lower bin limit of the DFO allele definitions for both the initial survey data and in the 192-fish blind sample test.

Results

The coordination and standardization process initially involved the creation of a shared tissue set representing selected populations from southern British Columbia and Washington. Advice was sought from colleagues in other laboratories concerning appropriate loci to consider for inclusion in the standardized baseline. Independent genotyping of the individual fish in the shared tissue set was conducted by both laboratories. Loci of potential value for inclusion in the standardized baseline were surveyed for chum salmon. A workshop was held where representatives from both laboratories exchanged and reviewed electropherograms, data, and lab-specific protocols for all proposed loci. The loci selected for further evaluation are outlined in Table 1. These 15 loci are proposed to form the core set for developing a standardized baseline. Criteria for selecting candidate loci included ease and reliability of genotyping, information content of each locus (numbers and frequency distributions of alleles), and ease/reliability of PCR amplification and scope for multiplexing (at sequencer and/or PCR stage). The scope of existing baselines was also a factor in selecting loci for the standardized baseline.

Independent genotyping of all candidate loci was conducted by both laboratories using a second shared tissue set consisting of samples from 192 fish representing other selected populations from British Columbia and Washington. An evaluation was then conducted on the concordance of allele genotyping between the two laboratories in both the initial set of fish used in the preliminary analysis of genetic variation in chum salmon and in the subsequent test sample of 192 fish. Concordance in allele identification between the laboratories was greater than 99.5% for all 15 loci selected for inclusion in a standardized baseline in both the set of fish initially surveyed to evaluate genetic variation, and in the second (blind) sample of fish specifically analyzed to test for concordance of allele identification between laboratories (Table 2). The agreement between laboratories in consistent allele identification over the complete size range of alleles observed at each locus is illustrated in Figures 1-15. The size range of alleles observed at each locus was substantial, and frequency of alleles in the fish surveyed displayed considerable variation. Despite variation in allele frequencies among loci, the concordance in identification of alleles between laboratories was consistent over the size range of alleles observed in all loci (Figures 16-30).

Discussion

Genetic methods of stock identification have several advantages over other techniques, among them the level of differentiation among populations and the stability of the genetic characters surveyed. Although initially allozymes proved successful in local applications (e.g. Beacham et al. 1987), DNA-level variation has been demonstrated to be effective in applications involving more complex mixtures of populations, allowing identification to the individual population in complex regional assemblages of sockeye salmon populations (Beacham et al. 2004, 2005). These genetic differences among populations are generally stable over time frames of interest in management applications.

The key requirement for application of any stock identification technique is accuracy of estimation of stock composition to the smallest practical unit, which in some cases can be to the local area, but in many cases requires identification to the river or lake of origin, and in the most demanding cases will require the identification of individual salmon to river or lake of origin. Microsatellites are useful in providing regional estimates of stock compositions, but they can also provide population-specific estimates in some applications if the survey of baseline populations has been adequate. As microsatellites have been demonstrated to be effective in stock identification applications in Pacific salmon, it seemed appropriate to develop a standardized microsatellite baseline that could be applied to estimate stock compositions in chum salmon fisheries in southern British Columbia and Washington. This project was thus developed as a cooperative effort between staff in the genetics laboratories of the Department of Fisheries and Oceans and the Washington Department of Fish and Wildlife. Estimation of microsatellite allele size at a locus will vary between labs dependent upon the actual primers used for PCR amplification and the equipment used to size-fractionate the amplified product. Therefore, it is necessary to standardize allele names among laboratories if a shared, standardized baseline is to be used in estimation of stock composition.

The results of the current study indicated agreement was reached between the Department of Fisheries of Oceans and the Washington Department of Fish and Wildlife pertaining to the microsatellite loci to be included in a shared baseline, and that consistent identification of alleles observed at these loci was obtained between staff in the two laboratories. Development of a standardized microsatellite baseline for stock

identification applications requires agreement as to the loci to include in the survey, and consistency in allele identification among laboratories. Both facets were achieved in the study, illustrating that development of a standardized microsatellite baseline is readily achievable. The next step in the application is to develop the standardized baseline, which will require surveying microsatellite variation in the 15 loci in selected chum salmon populations in southern British Columbia and Washington.

Literature Cited

Beacham, T.D., J.R. Candy, B. McIntosh, C. MacConnachie, A.Tabata, K. Kaukinen, L. Deng, K. M. Miller, R. E. Withler, and N. V. Varnavskaya. 2005. Estimation of stock composition and individual identification of sockeye salmon (*Oncorhynchus nerka*) on a Pacific Rim basis using microsatellite and major histocompatibility complex variation. Trans. Am. Fish. Soc. In press.

Beacham, T.D., A.P. Gould, R.E. Withler, C.B. Murray and L.W. Barner. 1987. Biochemical genetic survey and stock identification of chum salmon (Oncorhynchus keta) in British Columbia. Can. J. Fish. Aquat. Sci. 44: 1702-1713.

Beacham, T.D., M. Lapointe, J.R. Candy, B. McIntosh, C. MacConnachie, A. Tabata, K. Kaukinen, L. Deng, K.M. Miller, and R.E. Withler. 2004. Stock identification of Fraser River sockeye salmon (*Oncorhynchus nerka)* using microsatellites and major histocompatibility complex variation. Trans. Am. Fish. Soc. 133: 1106-1126.

Table 1. Summary of characteristics of loci surveyed in chum salmon. Range in allele size is listed in bp. He is expected heterozygosity, and Ho is observed heterozygosity.

Locus	Rank	Tag	PCR Reliability WDFW/DFO	Bins(bp)	Range	No. alleles	He	Ho	Comments (scoring)
One111	1	6FAM	2/3	2,4	140-500+	95	0.91	0.88	Good, clean
One102	1	NED	2/3	4	122-380	53	0.90	0.87	Good, clean
Ots3	1	HEX	2/1	2	61-150	31	0.73	0.72	Good with artifacts (common signals=104.5, 108.7bp)
One101	1	HEX	2/2	4	105-305	46	0.86	0.83	Good, clean
Ssa419	1	6FAM	2/3	4	180-450	37	0.81	0.81	Good, clean
Oki2	2	6FAM	2/3	4	117-290	31	0.83	0.83	Some individuals with 3or4 alleles, A+splits
One103	1.5	HEX	2/2	4	90-325	48	0.87	0.83	Good with artifacts (common signals=94.7, 98.7, 118.5, 122.6, 126.5, 130.5bp)
One104	1	NED	2/3	4	102-276	33	0.91	0.90	Good, A+splits
Oke3	1	NED	2/2	13	204-373	13	0.64	0.63	Good with minor artifacts (=mirror image peaks)
OtsG68	1	6FAM	2/2	4	132-450	51	0.92	0.92	Good, clean
Omy1011	1	HEX	2/1	4	162-320	37	0.89	0.88	Good, A+splits
Oki100	1	NED	2/2	4	71-192	25	0.80	0.80	Good, clean
One114	1	NED	2/2	4	164-430	49	0.90	0.88	Good, clean
Omm1070	1	6FAM	3	4	222-353	33	0.96	0.91	Good, clean signal
Ots103	1	NED	2	4	88-249	39	0.95	0.95	Good, clean signal

	Overall ranking	PCR Reliability
1	Desirable	<5% failures
2	Acceptable	5%< failures <15%
3	Undesirable	>15% failures

Table 2: Total number of allele comparisons and percentage concordance at designated alleles estimated by DFO and WDFW laboratories for all microsatellite loci in standardized and blind sample sets.

		Percentage Concordance		
Locus	Total Number of Comparisons	Standardized sample set	Blind sample set	Mean
One111	634	99.70%	100.00%	99.84%
One102	634	100.00%	100.00%	100.00%
Ots3	638	100.00%	100.00%	100.00%
One101	616	99.69%	99.66%	99.68%
Ssa419	638	99.70%	100.00%	99.84%
Oki2	632	100.00%	100.00%	100.00%
One103	620	100.00%	99.66%	99.84%
One104	628	100.00%	100.00%	100.00%
Oke3	664	99.72%	100.00%	99.85%
Otsg68	618	100.00%	100.00%	100.00%
Omy1011	652	100.00%	99.33%	99.69%
Oki100	650	100.00%	100.00%	100.00%
One114	636	100.00%	100.00%	100.00%
Omm1070	568	100.00%	100.00%	100.00%
Ots103	630	100.00%	99.68%	99.84%

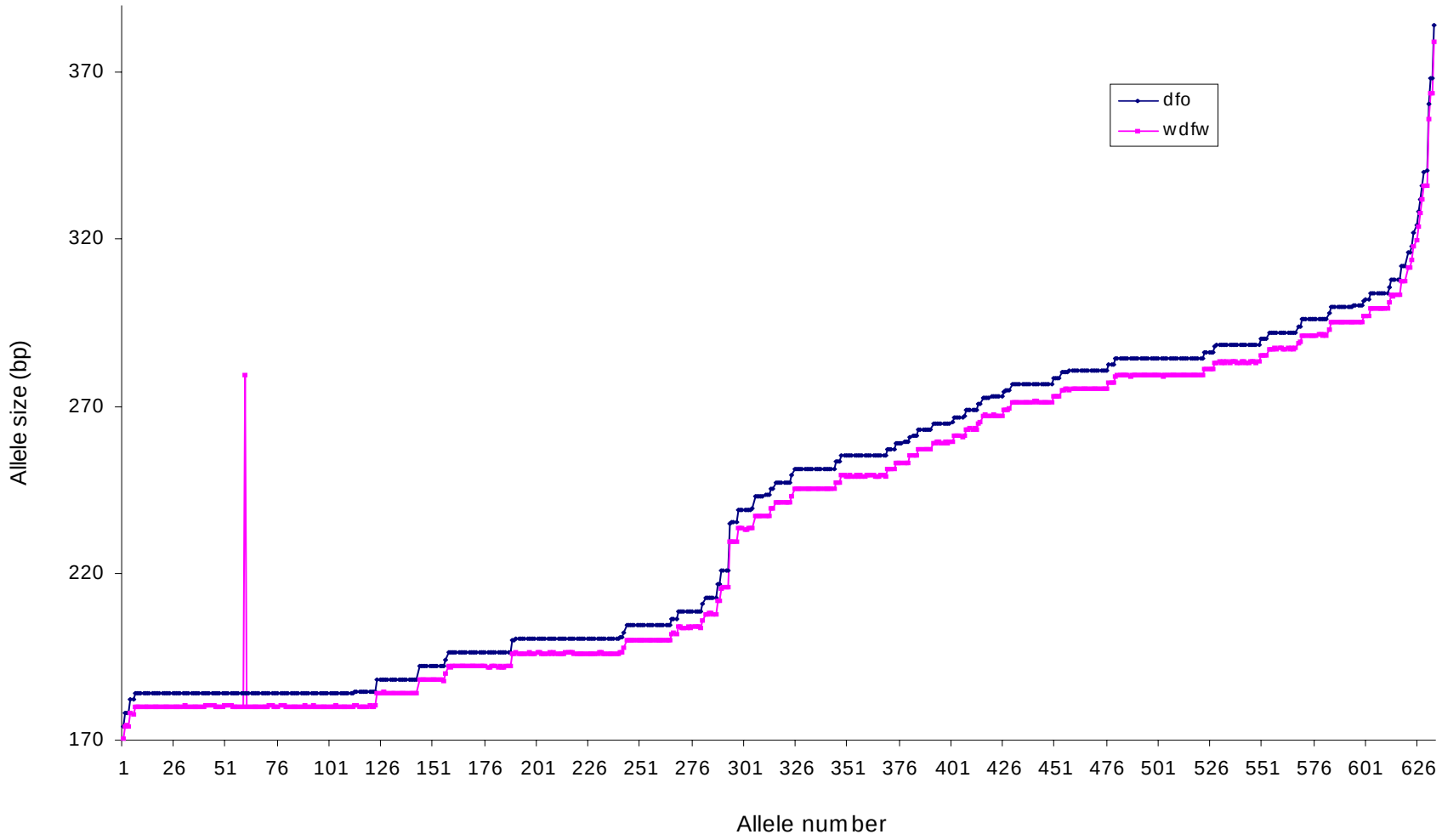

Figure 1: Concordance of allele size at One111 estimated by DFO and WDFW laboratories.

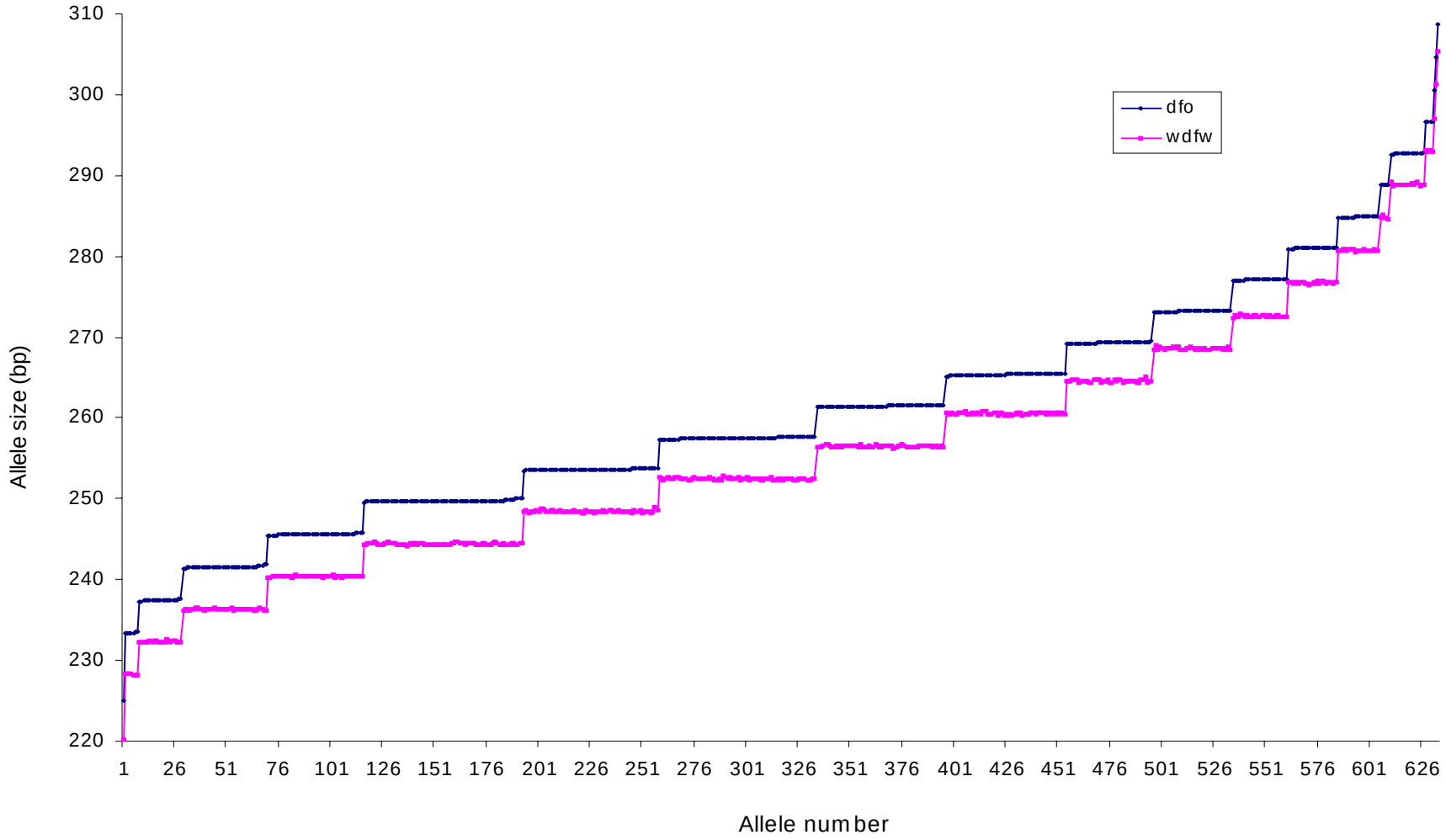

Figure 2: Concordance of allele size at One102 estimated by DFO and WDFW laboratories.

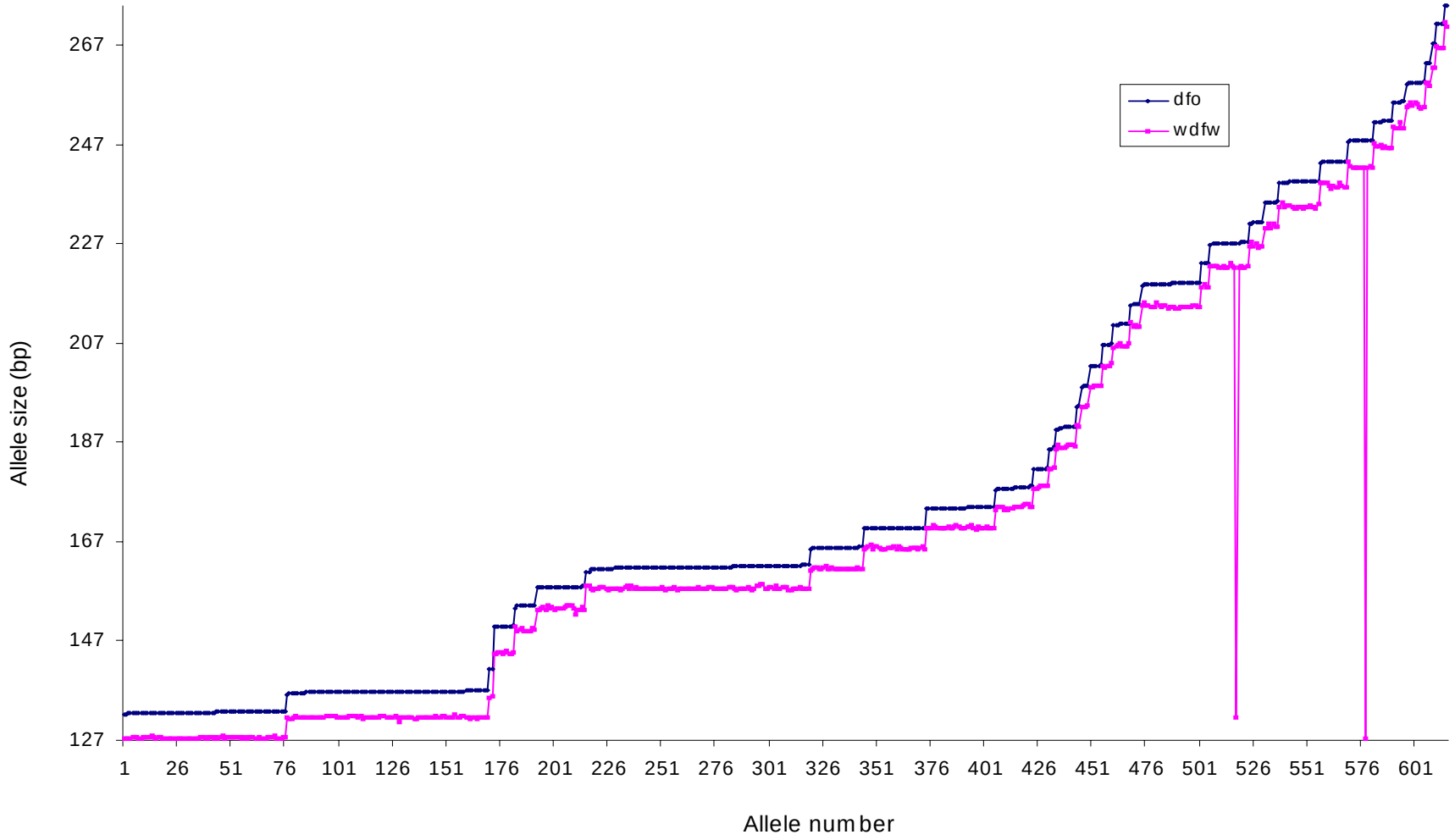

Figure 3: Concordance of allele size at One101 estimated by DFO and WDFW laboratories.

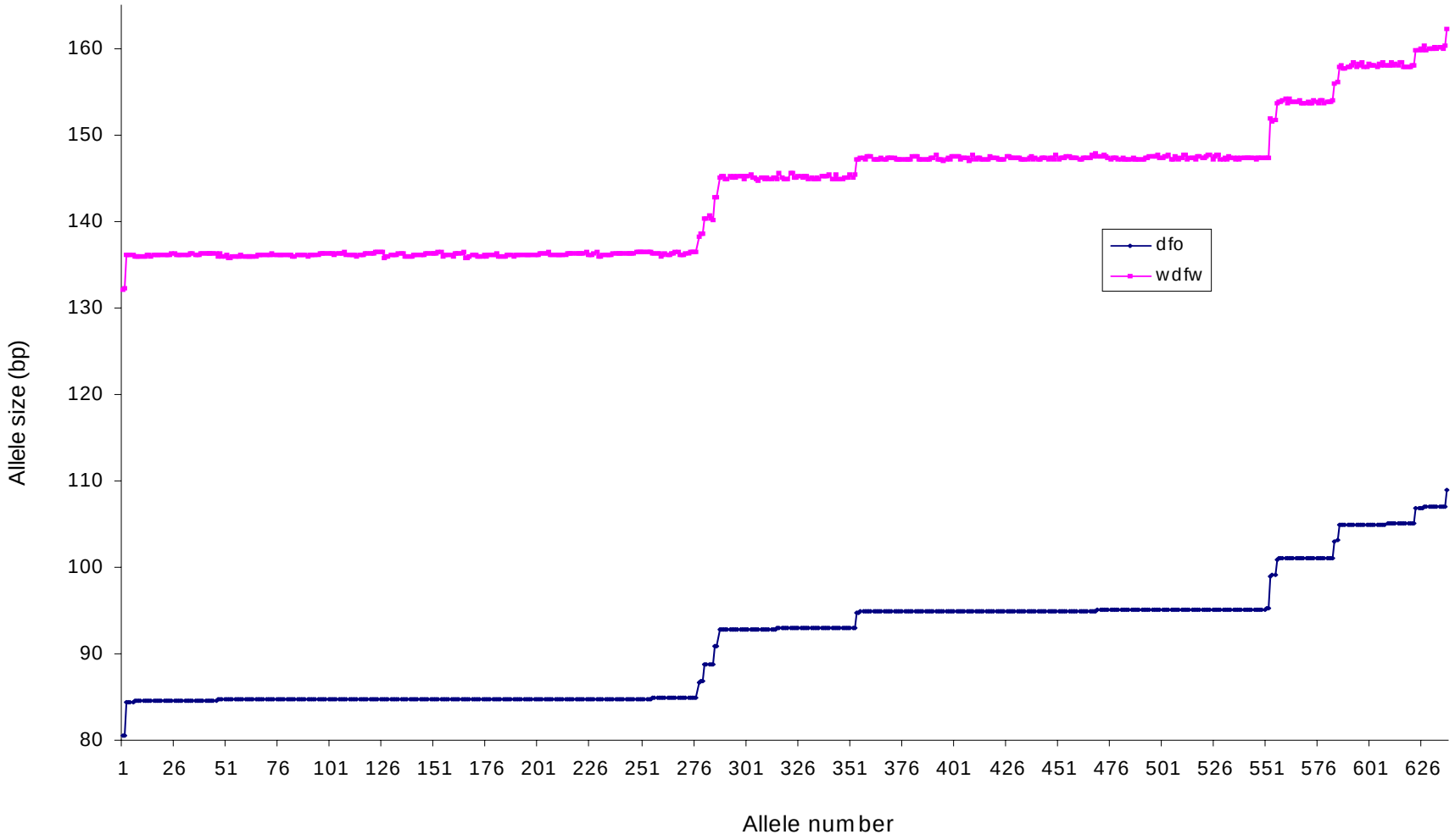

Figure 4: Concordance of allele size at Ots3 estimated by DFO and WDFW laboratories.

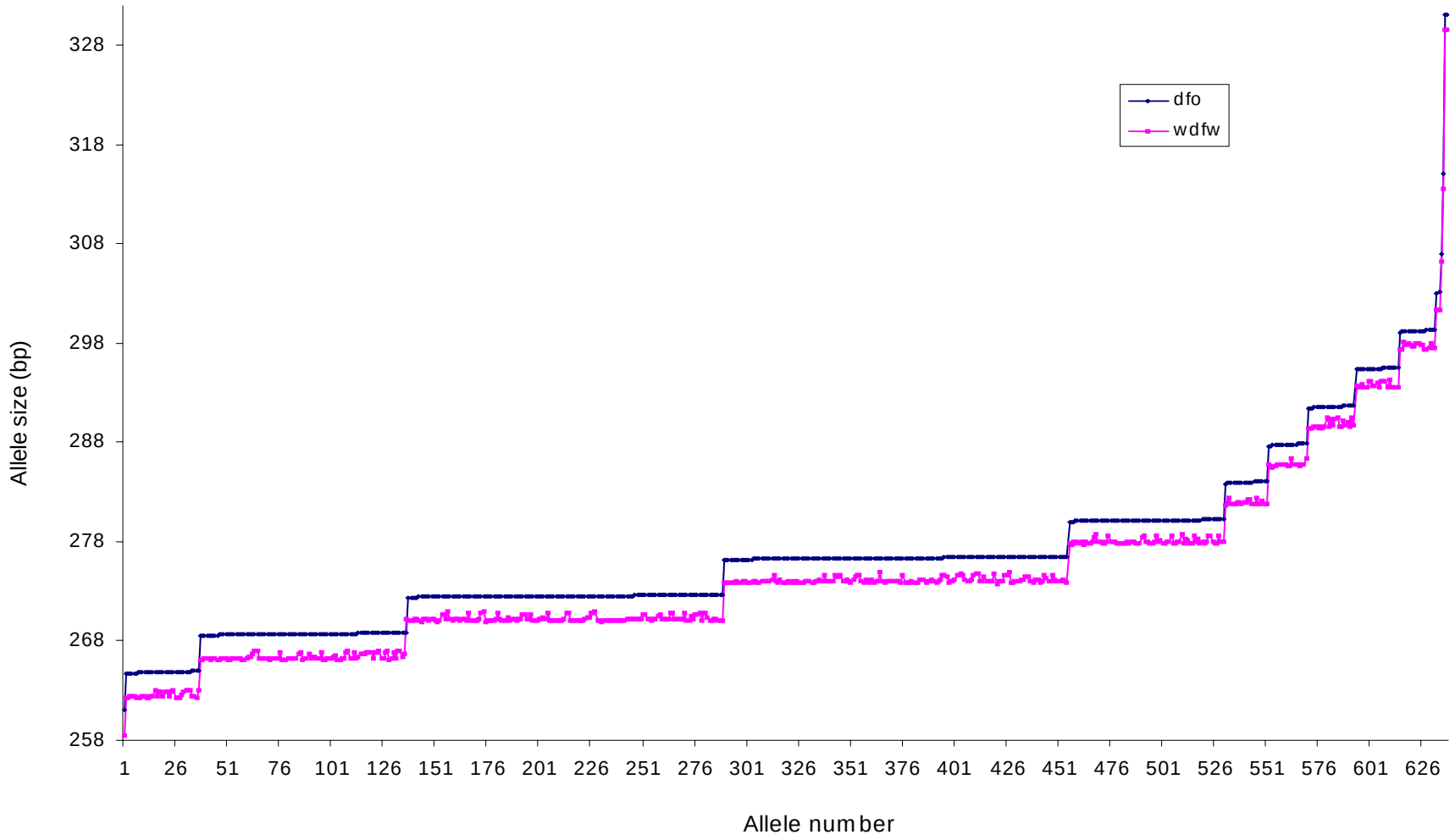

Figure 5: Concordance of allele size at Ssa419 estimated by DFO and WDFW laboratories.

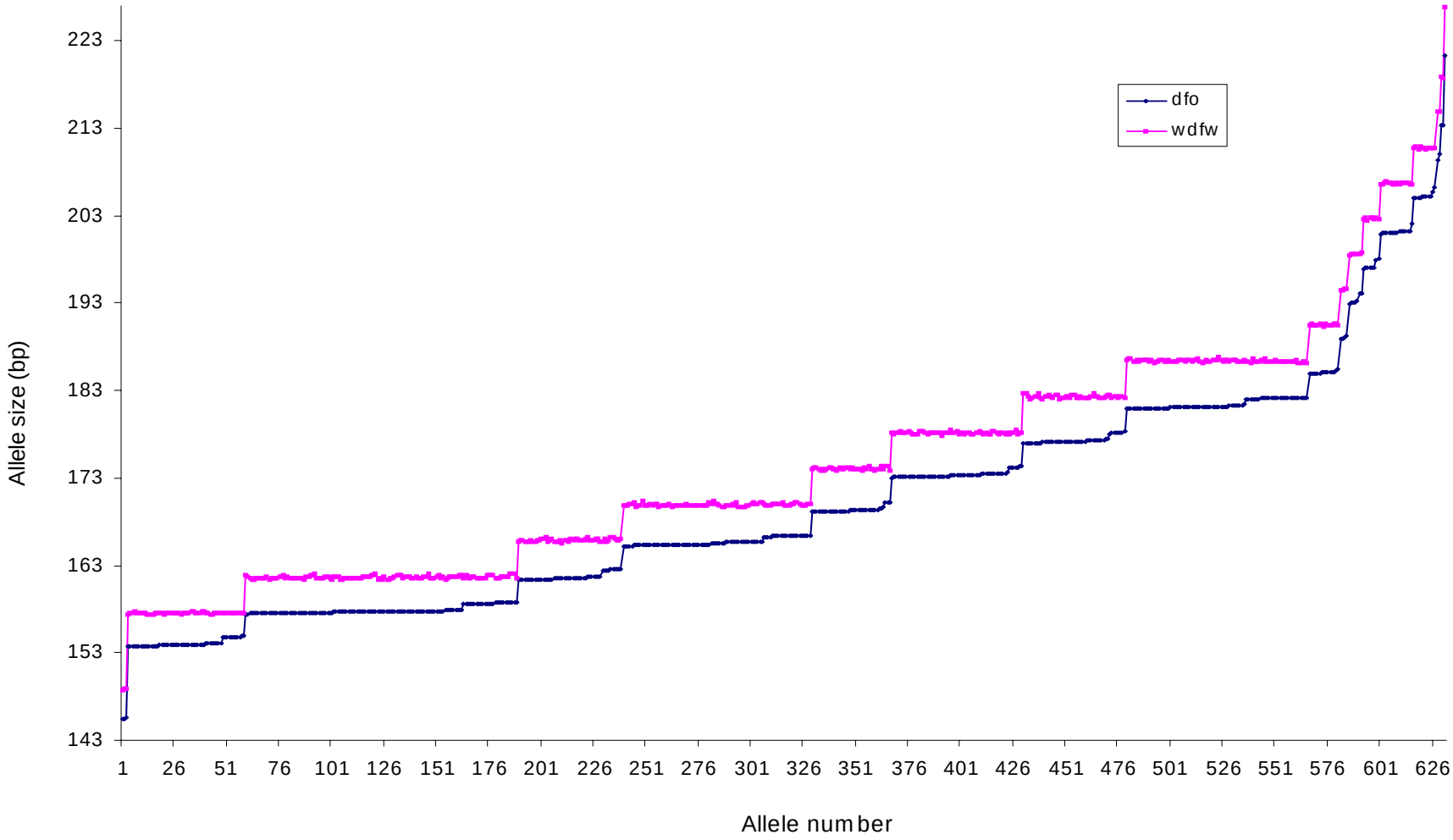

Figure 6: Concordance of allele size at Oki2 estimated by DFO and WDFW laboratories.

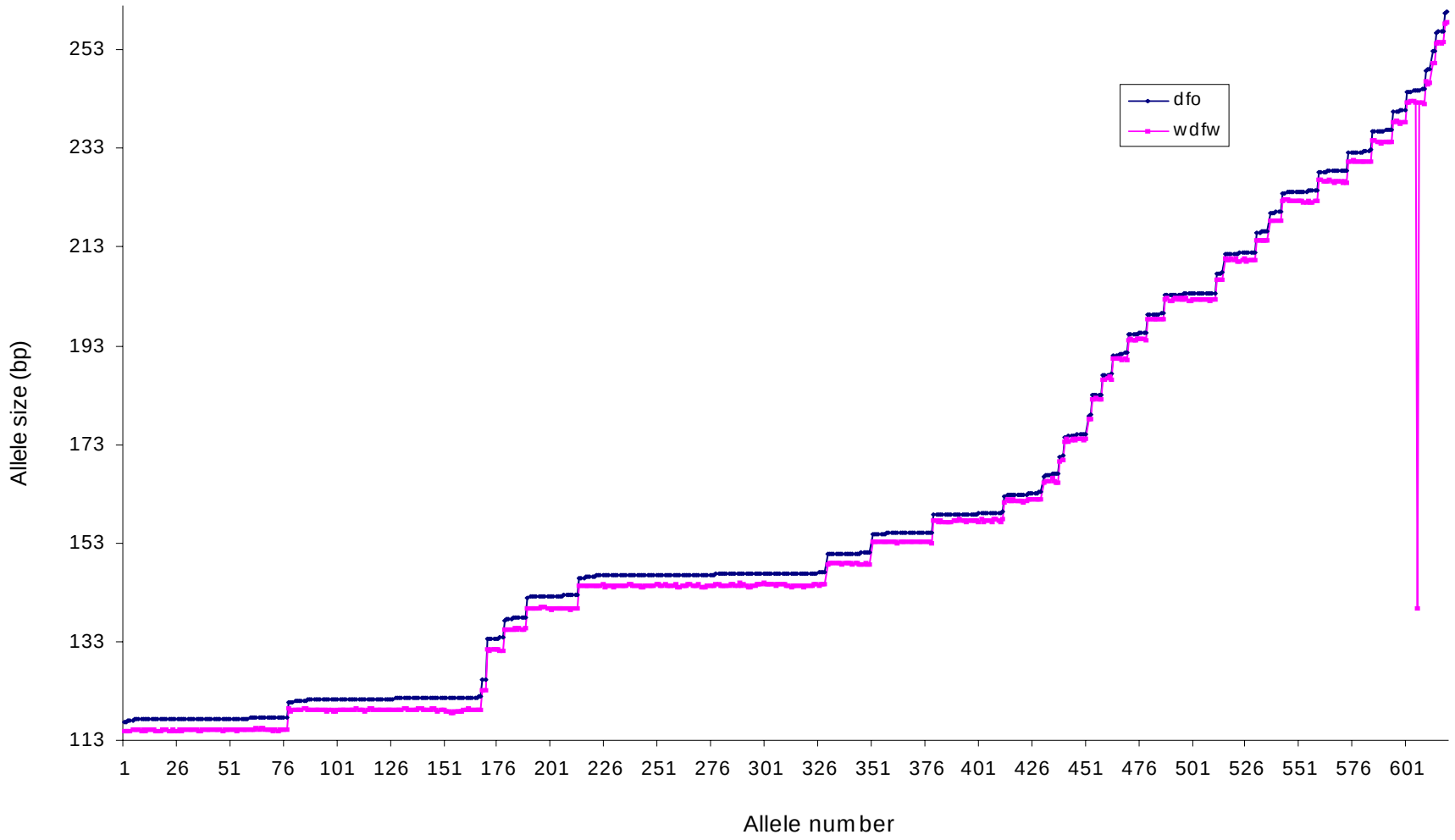

Figure 7: Concordance of allele size at One103 estimated by DFO and WDFW laboratories.

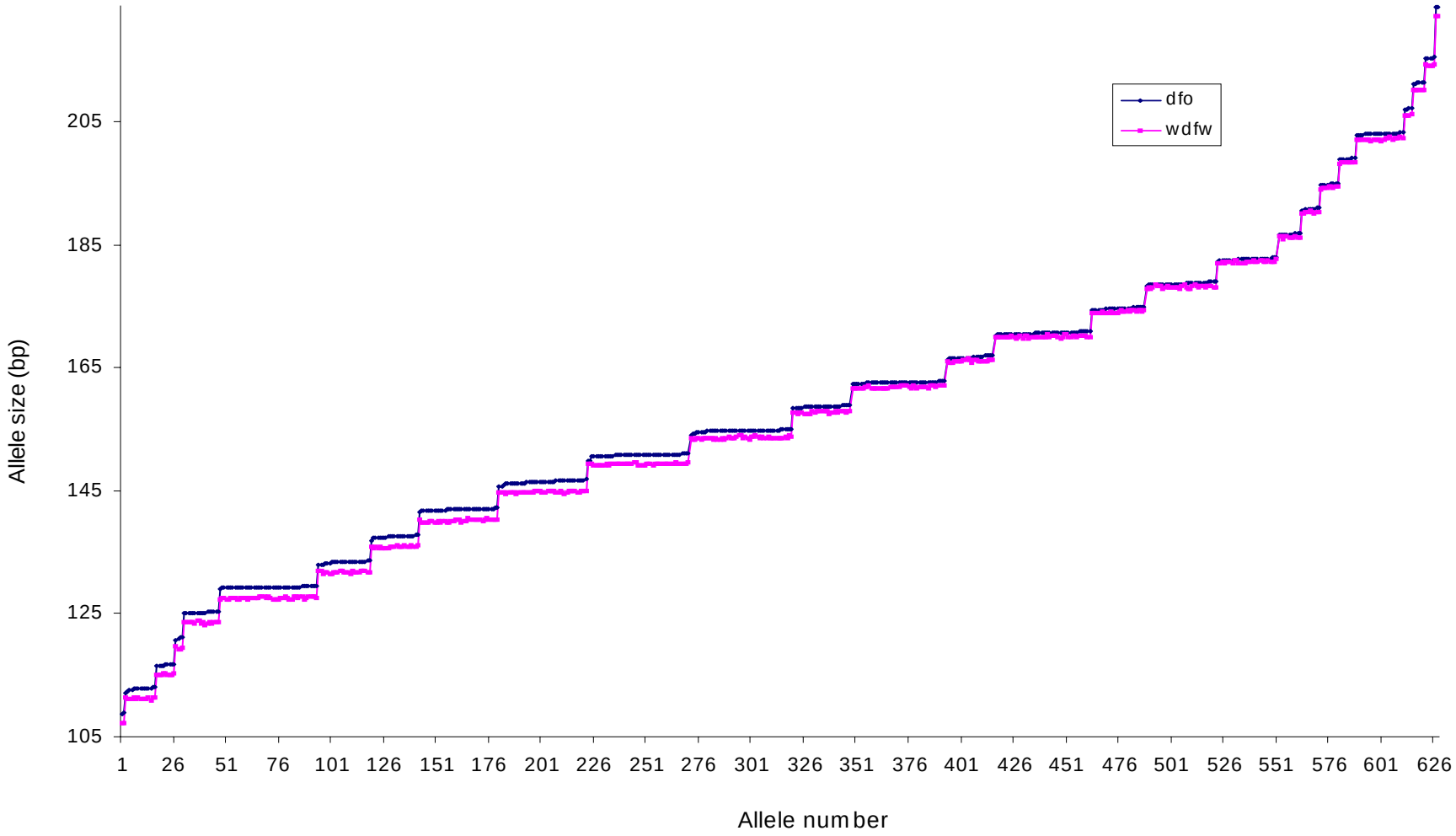

Figure 8: Concordance of allele size at One104 estimated by DFO and WDFW laboratories.

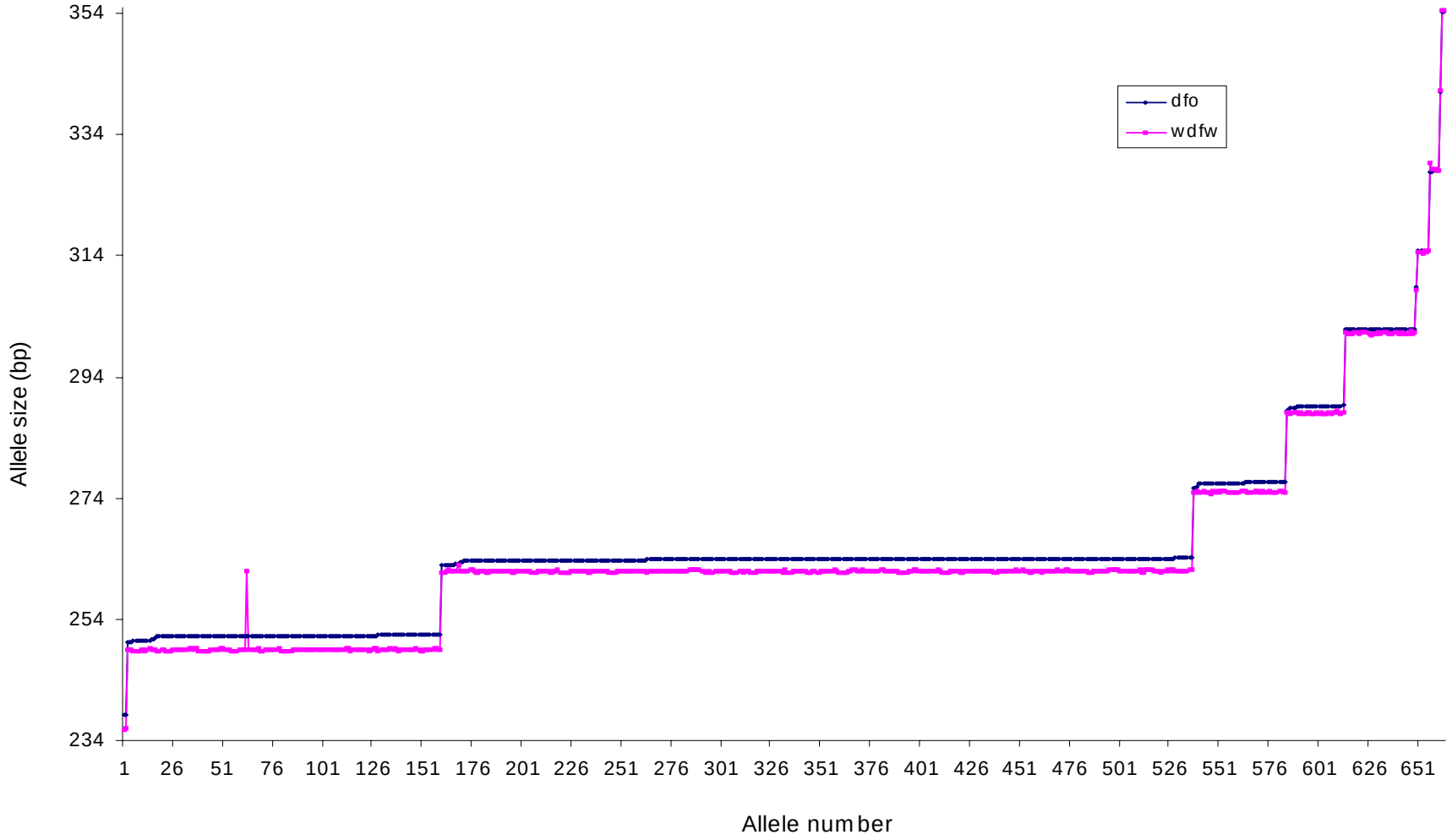

Figure 9: Concordance of allele size at Oke3 estimated by DFO and WDFW laboratories.

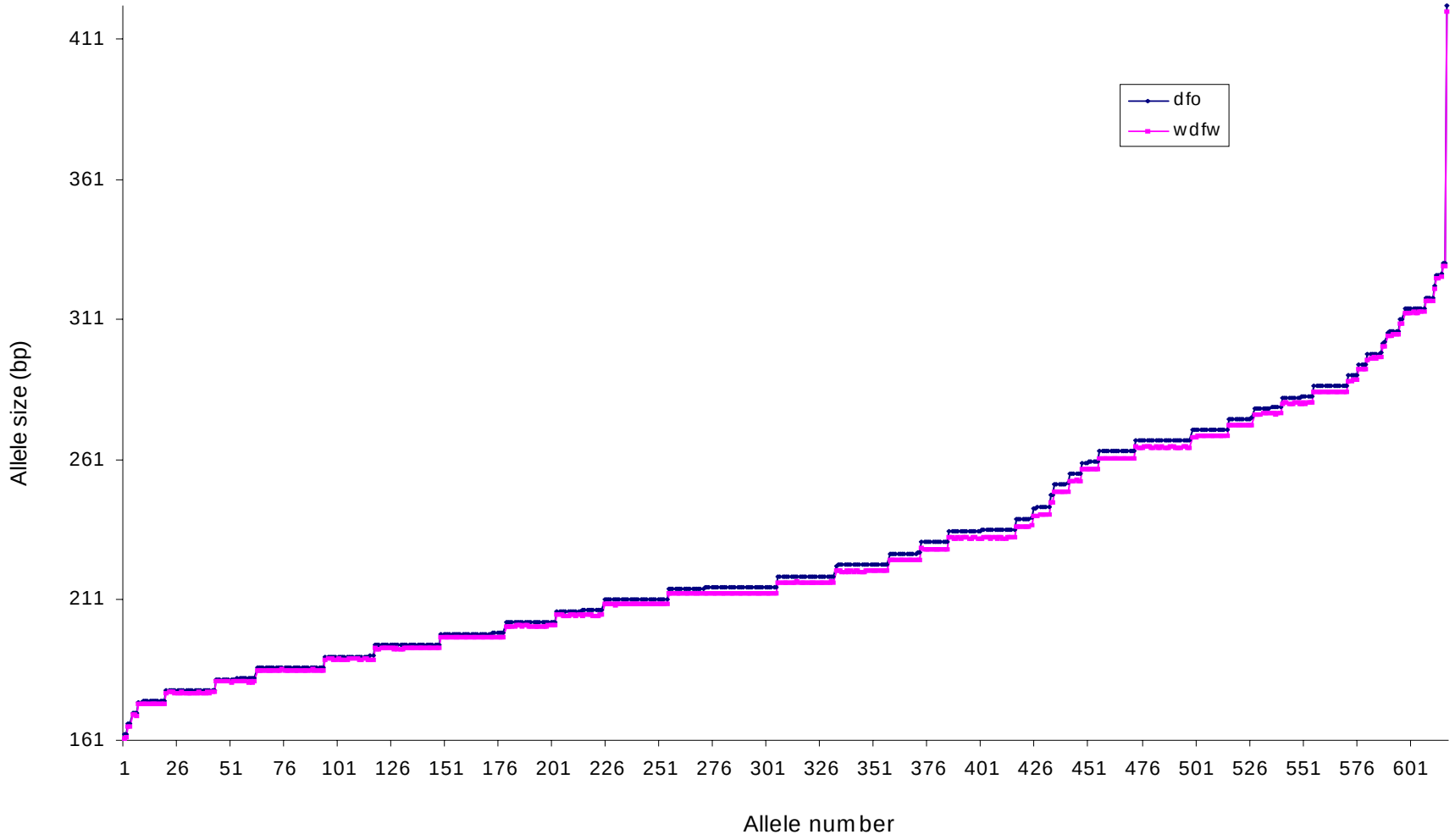

Figure 10: Concordance of allele size at Otsg68 estimated by DFO and WDFW laboratories.

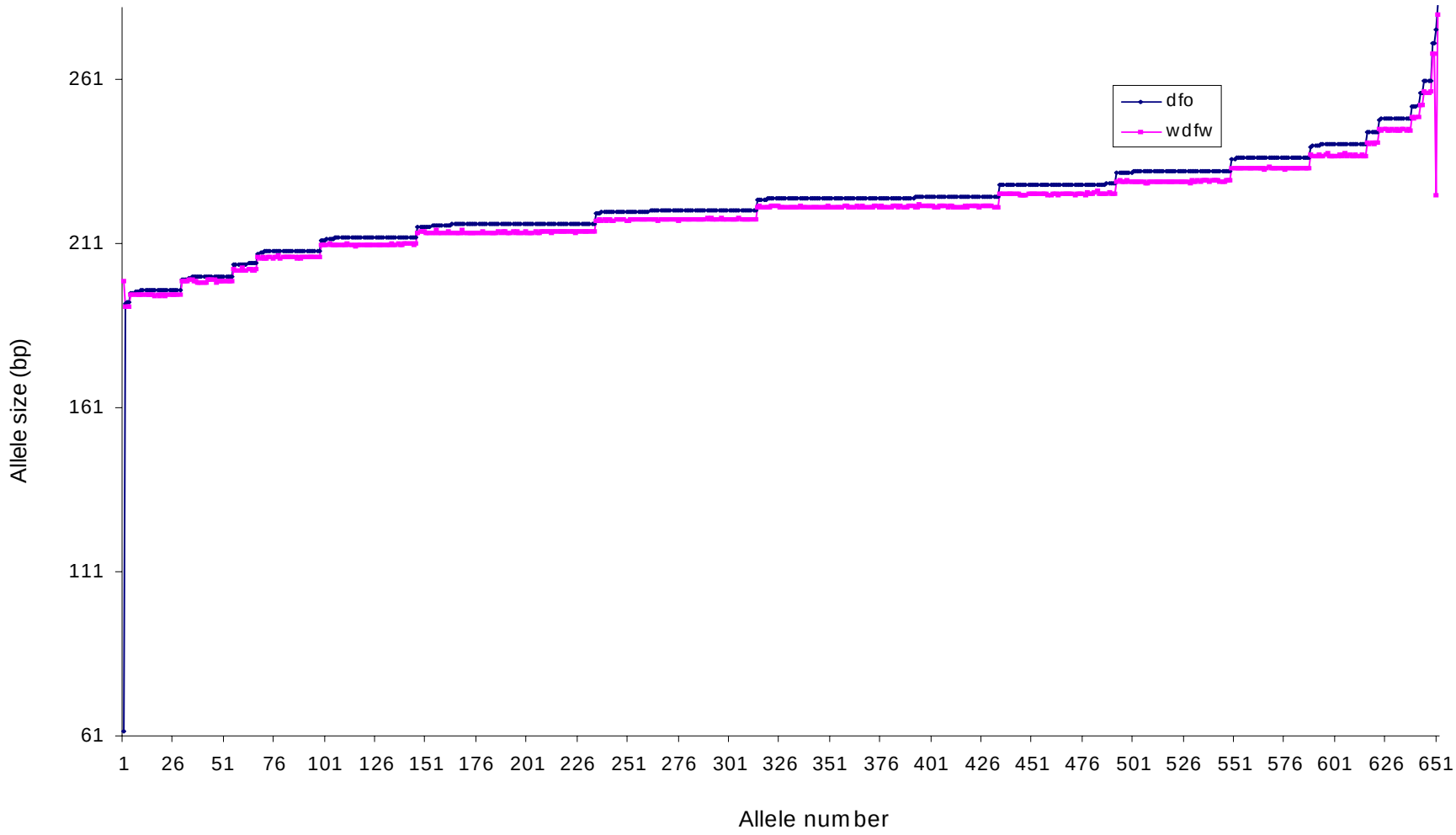

Figure 11: Concordance of allele size at Omy1011 estimated by DFO and WDFW laboratories.

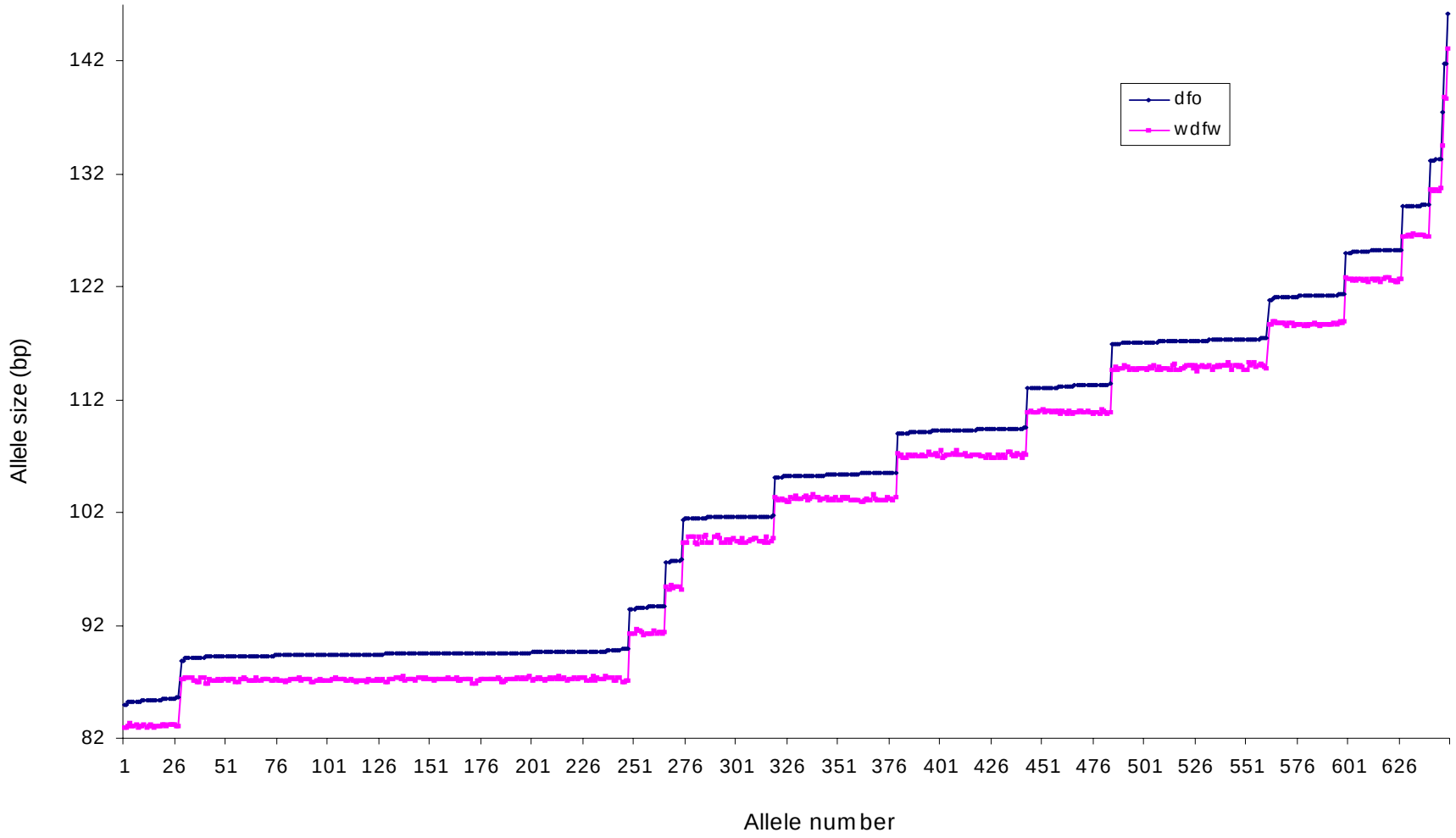

Figure 12: Concordance of allele size at Oki100 estimated by DFO and WDFW laboratories.

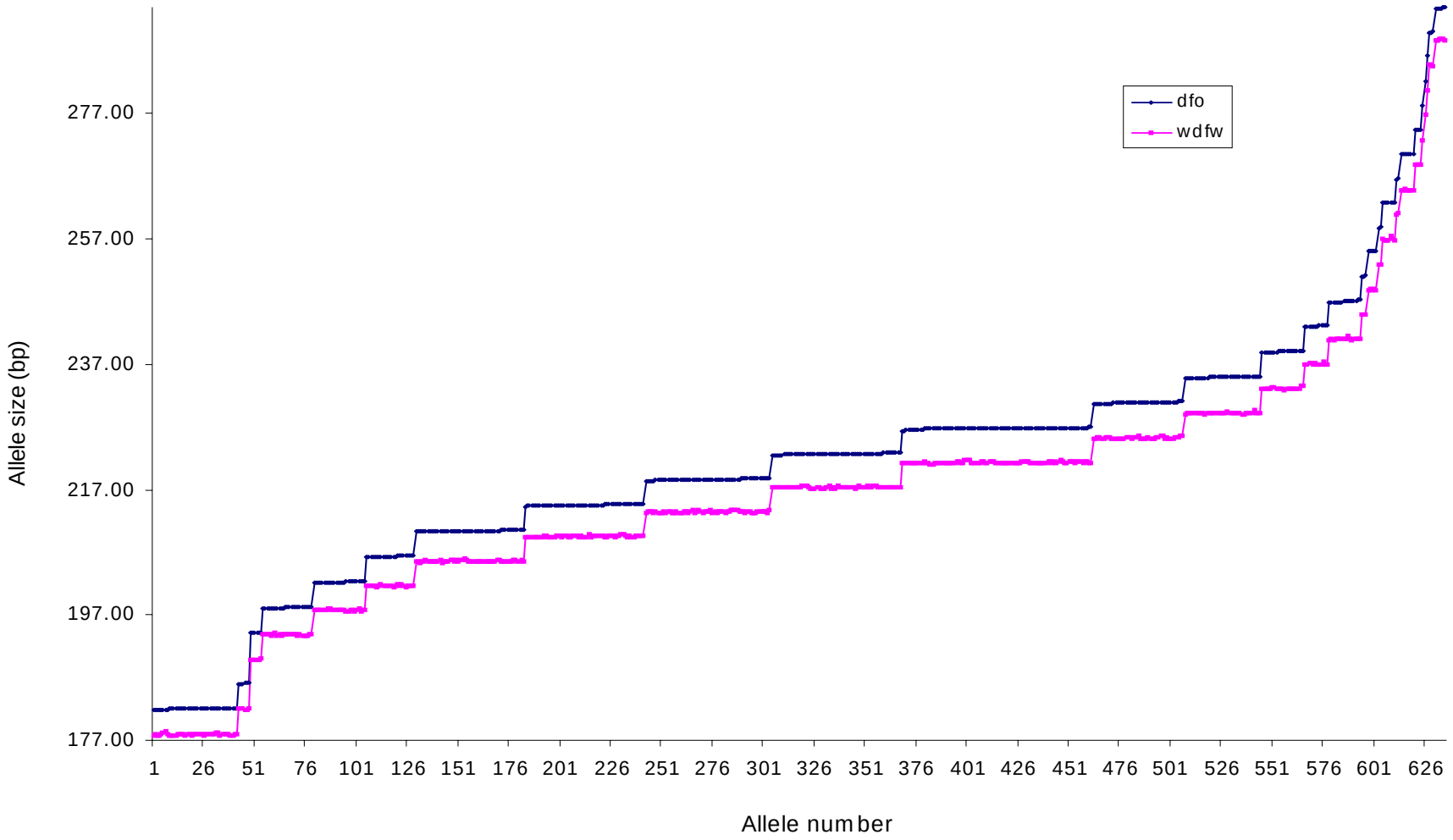

Figure 13: Concordance of allele size at One114 estimated by DFO and WDFW laboratories.

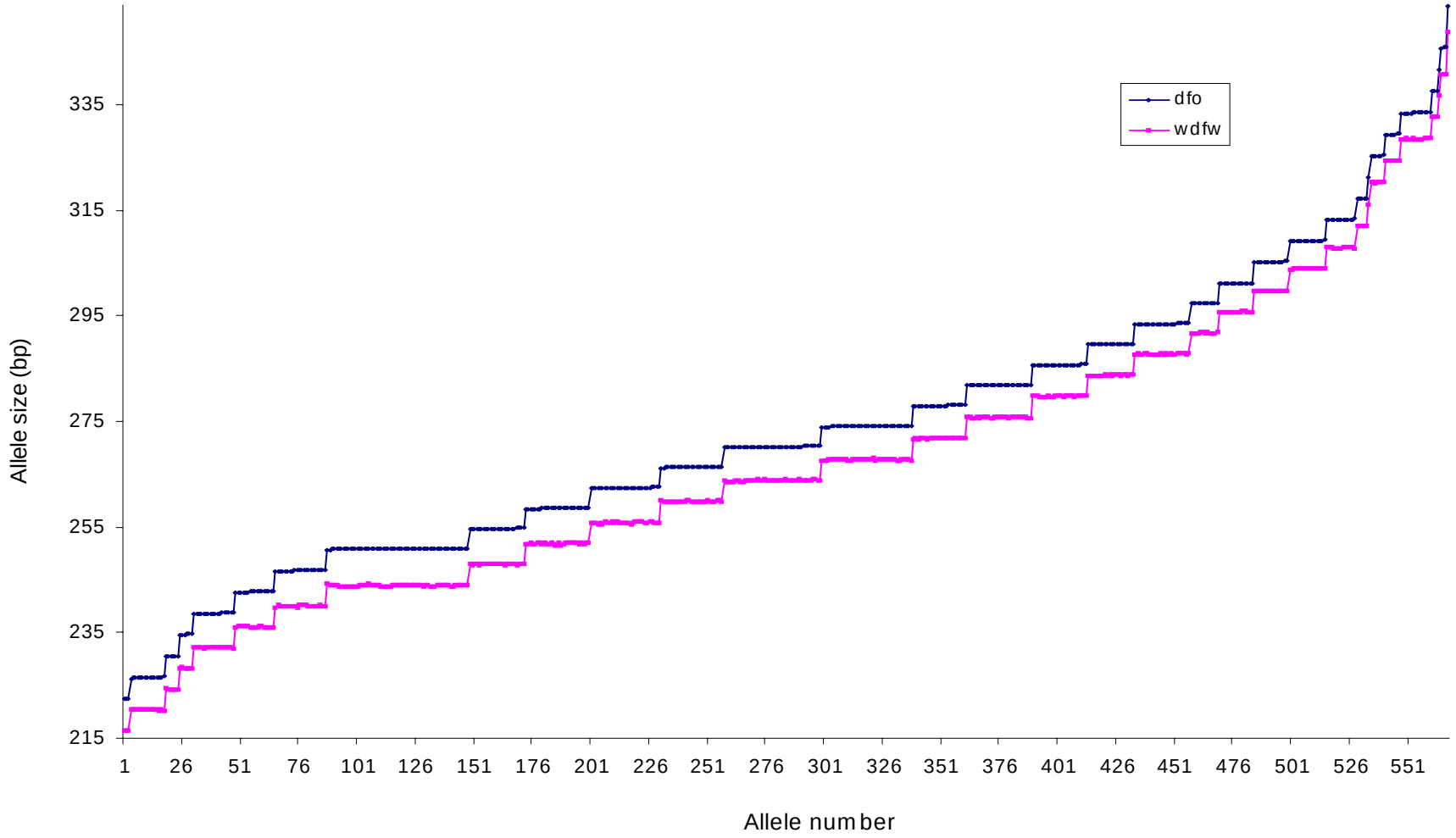

Figure 14: Concordance of allele size at Omm1070 estimated by DFO and WDFW laboratories.

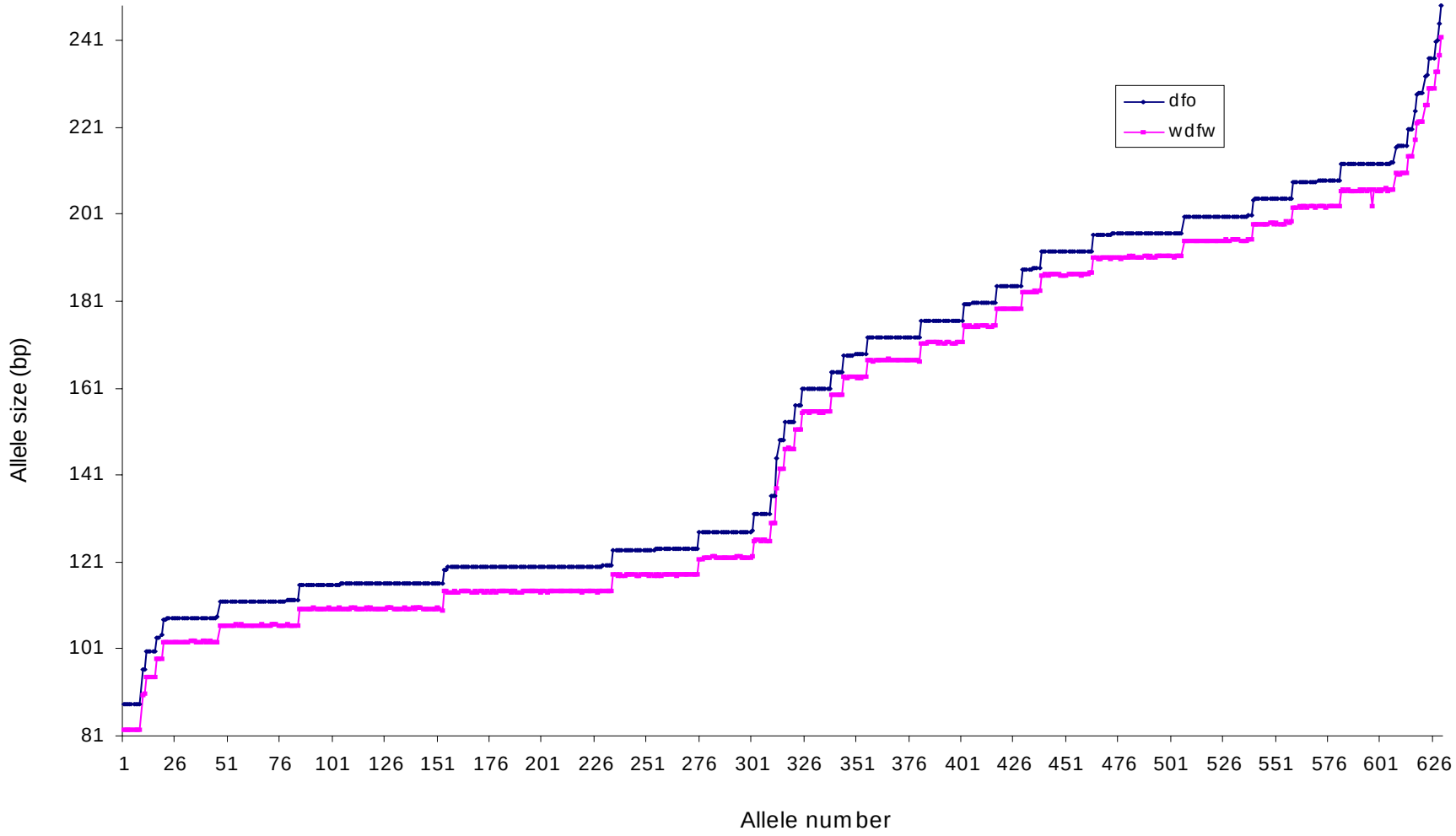

Figure 15: Concordance of allele size at Ots103 estimated by DFO and WDFW laboratories.

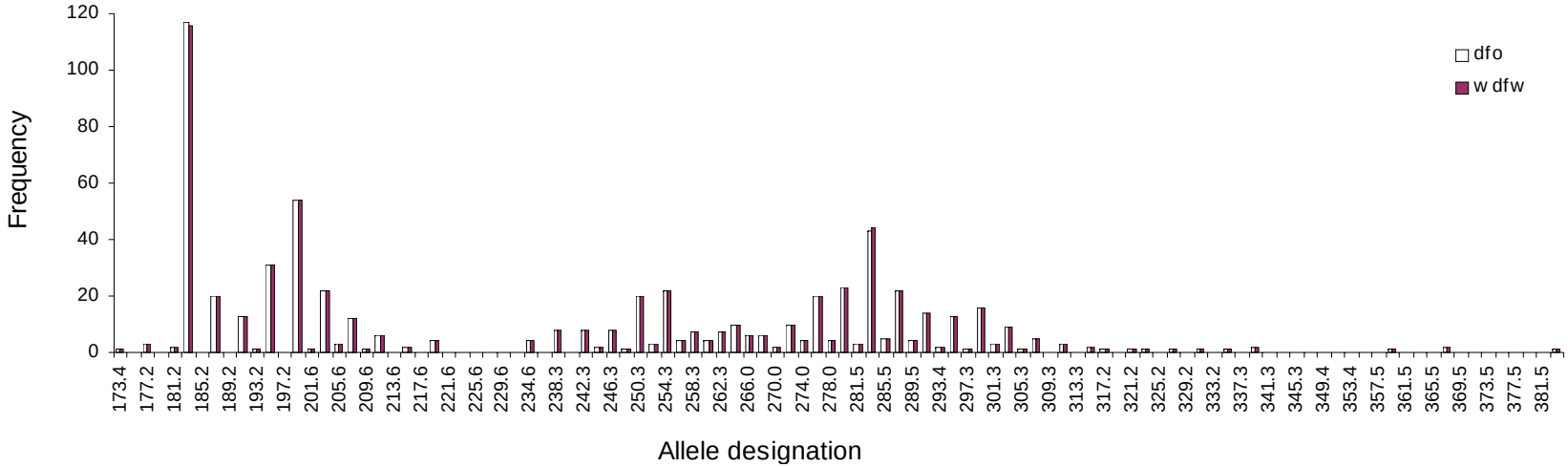

Figure 16: Concordance of designated alleles estimated by DFO and WDFW laboratories at One111 in standardized and blind sample sets.

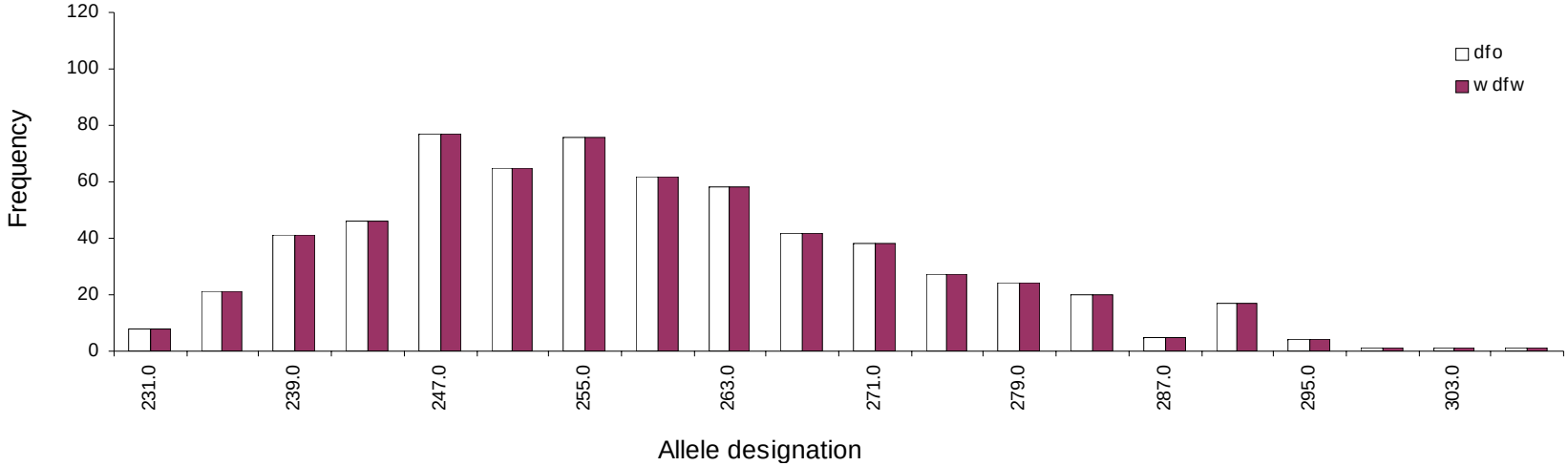

Figure 17: Concordance of designated alleles estimated by DFO and WDFW laboratories at One102 in standardized and blind sample sets.

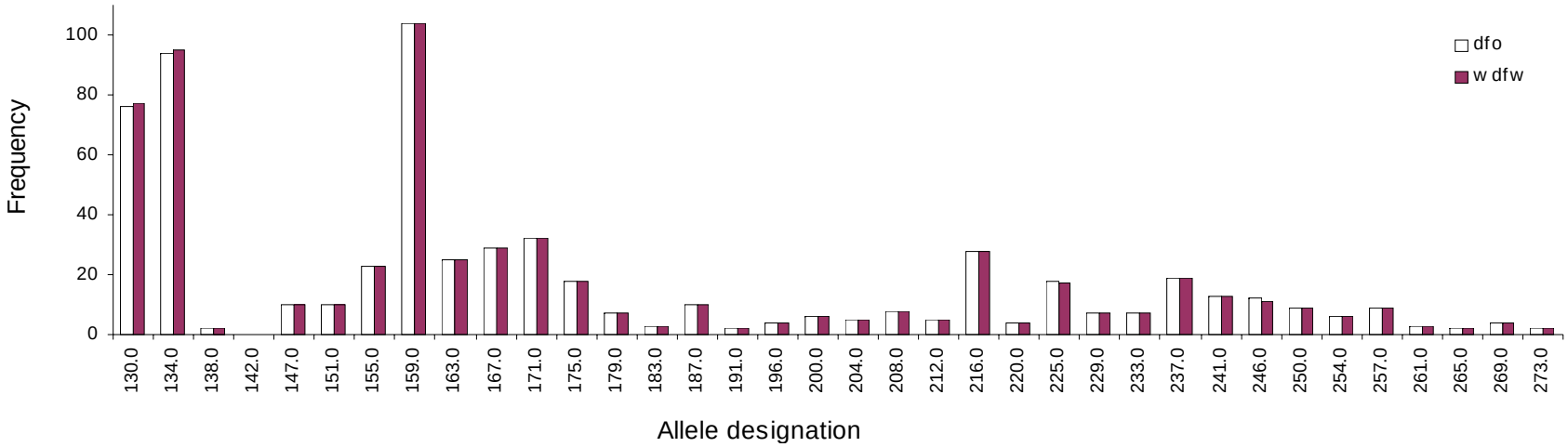

Figure 18: Concordance of designated alleles estimated by DFO and WDFW laboratories at One101 in standardized and blind sample sets.

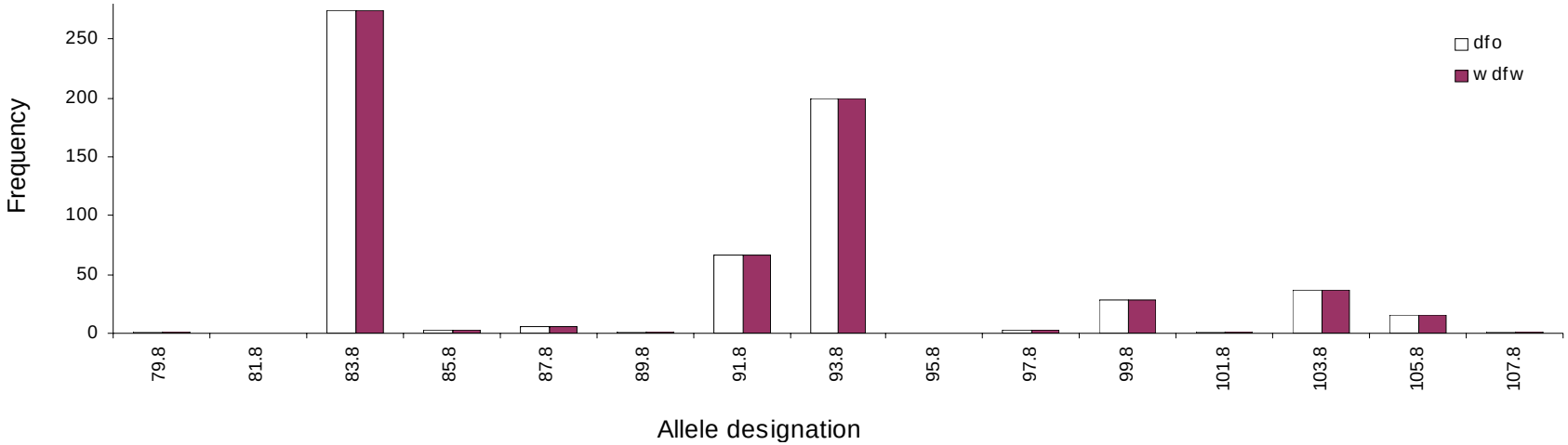

Figure 19: Concordance of designated alleles estimated by DFO and WDFW laboratories at Ots3 in standardized and blind sample sets.

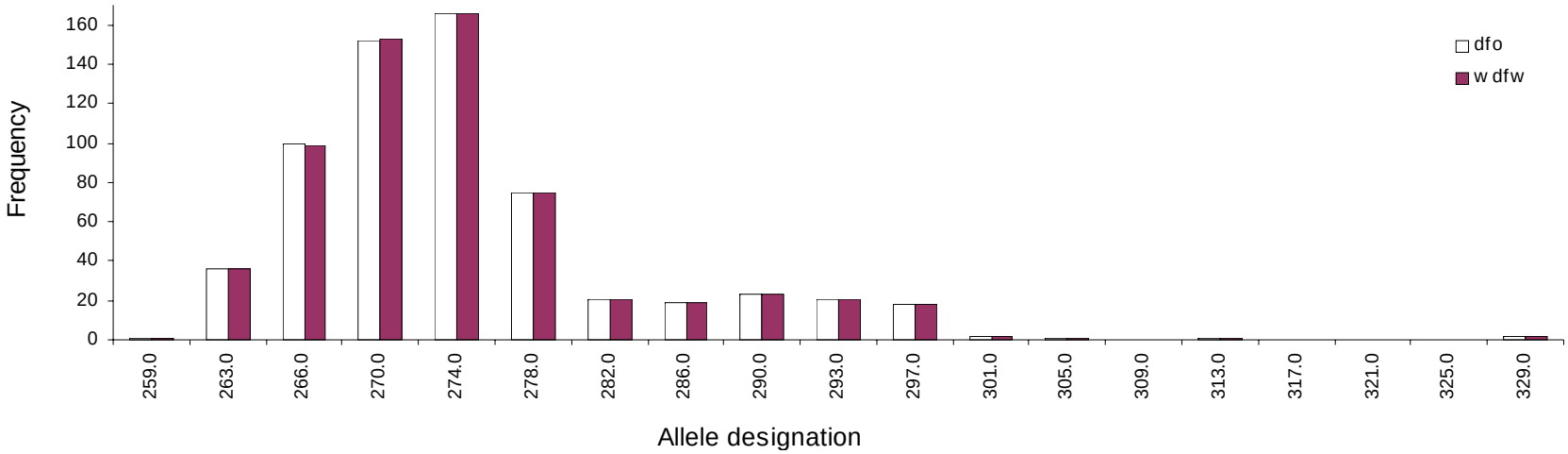

Figure 20: Concordance of designated alleles estimated by DFO and WDFW laboratories at Ssa419 in standardized and blind sample sets.

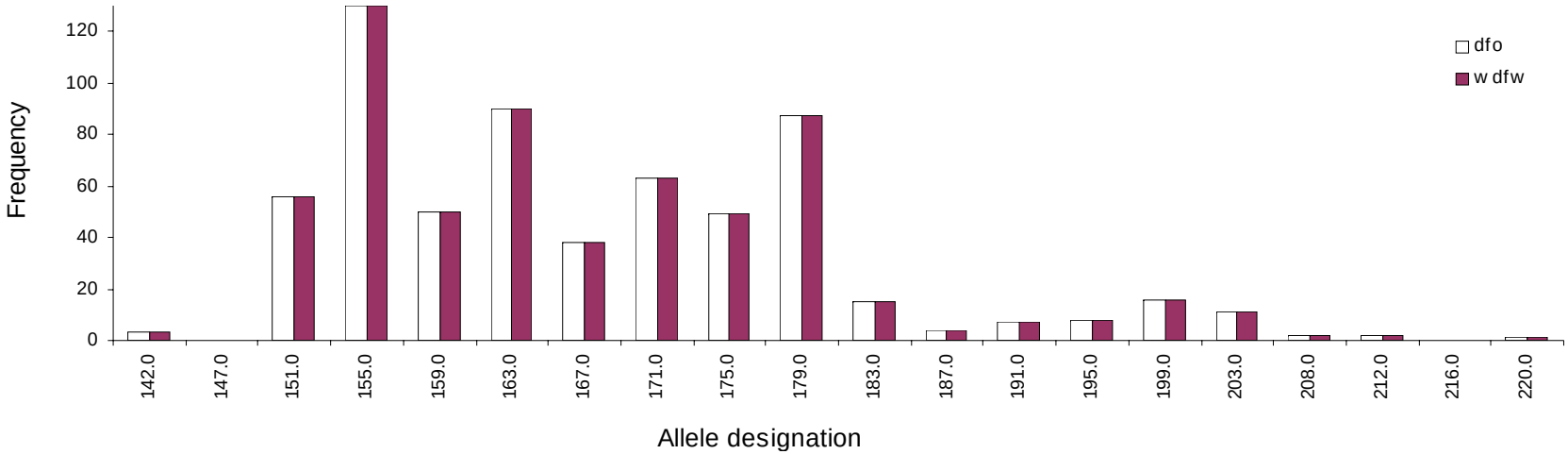

Figure 21: Concordance of designated alleles estimated by DFO and WDFW laboratories at Oki2 in standardized and blind sample sets.

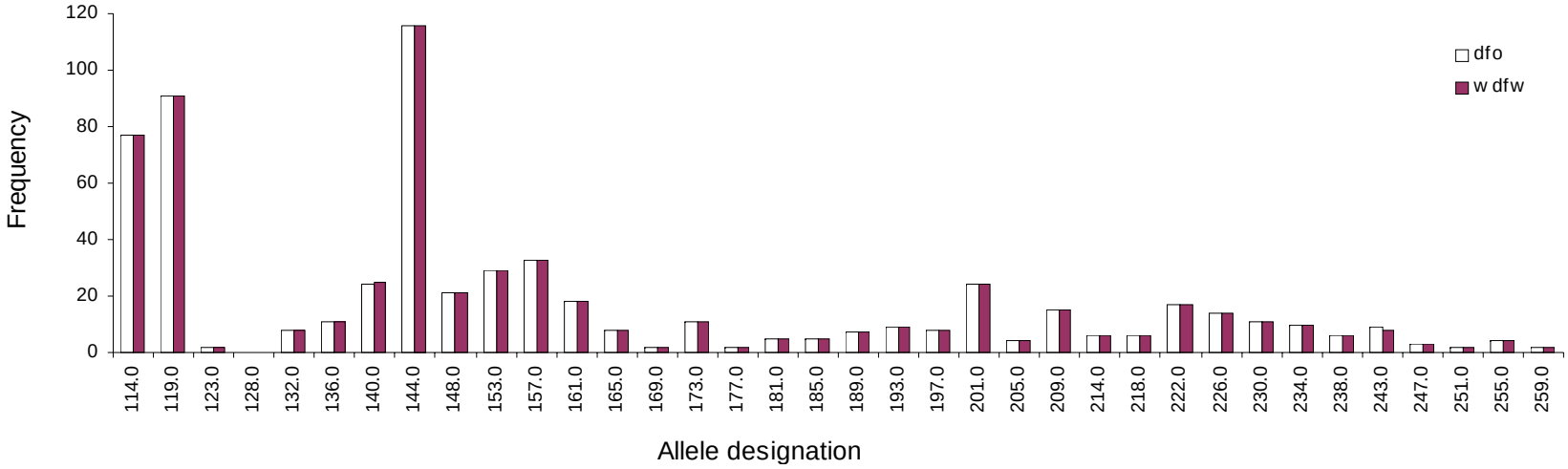

Figure 22: Concordance of designated alleles estimated by DFO and WDFW laboratories at One103 in standardized and blind sample sets.

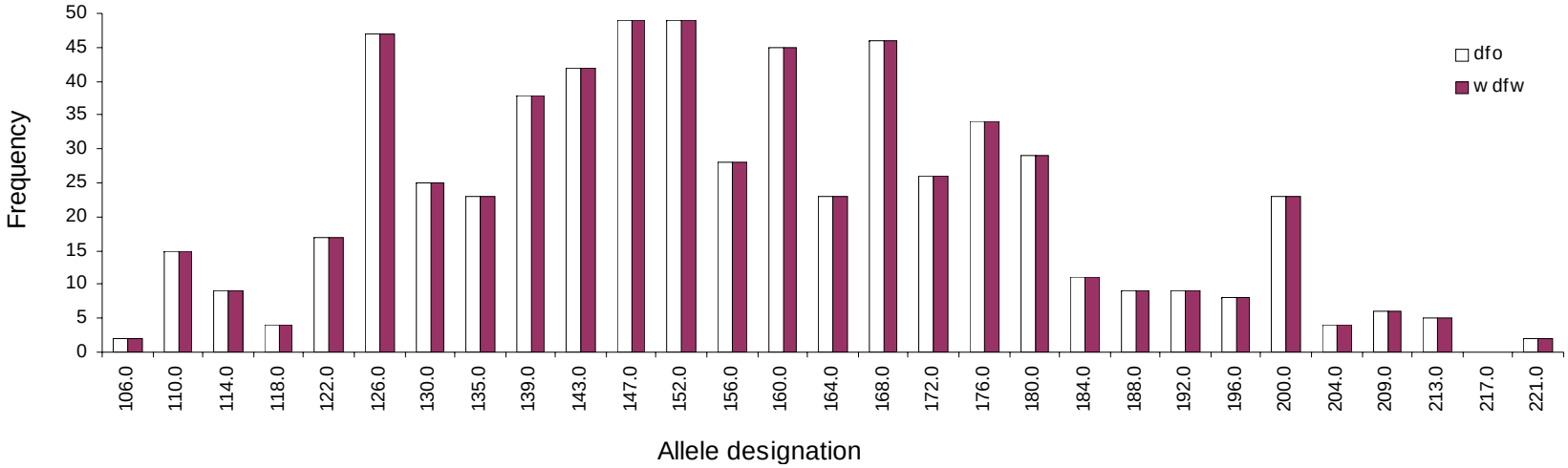

Figure 23: Concordance of designated alleles estimated by DFO and WDFW laboratories at One104 in standardized and blind sample sets.

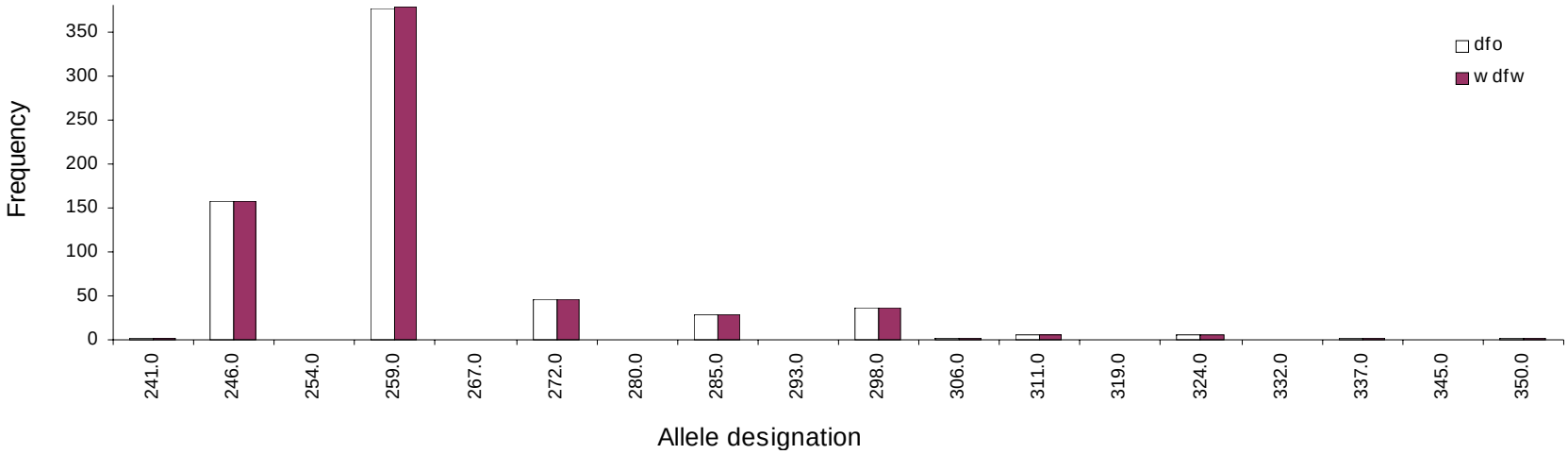

Figure 24: Concordance of designated alleles estimated by DFO and WDFW laboratories at Oke3 in standardized and blind sample sets.

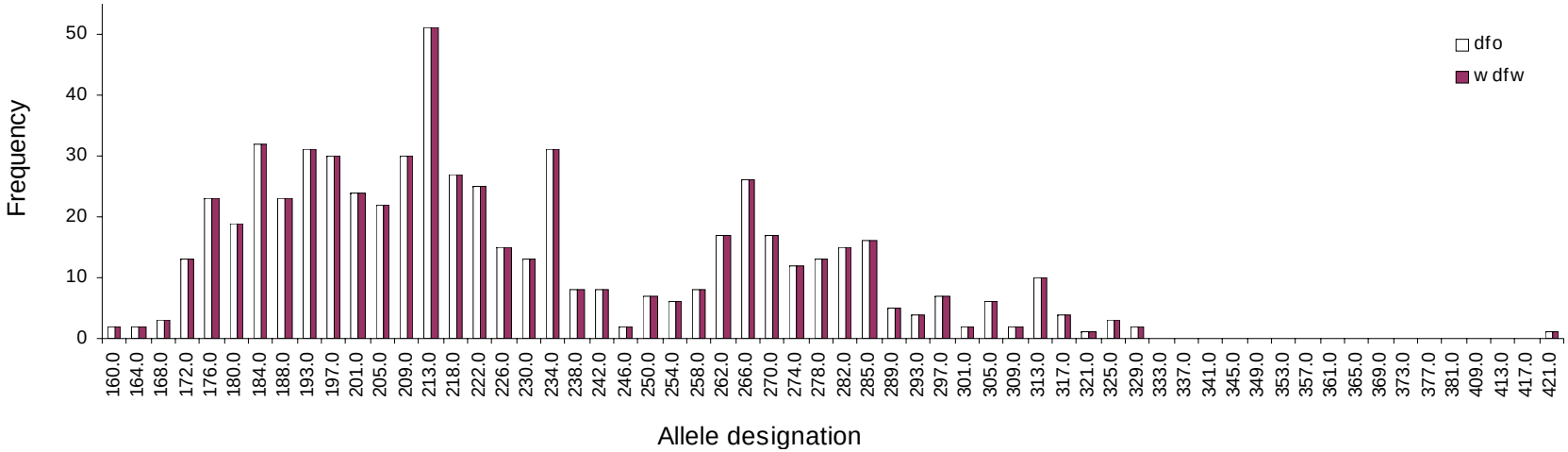

Figure 25: Concordance of designated alleles estimated by DFO and WDFW laboratories at Otsg68 in standardized and blind sample sets.

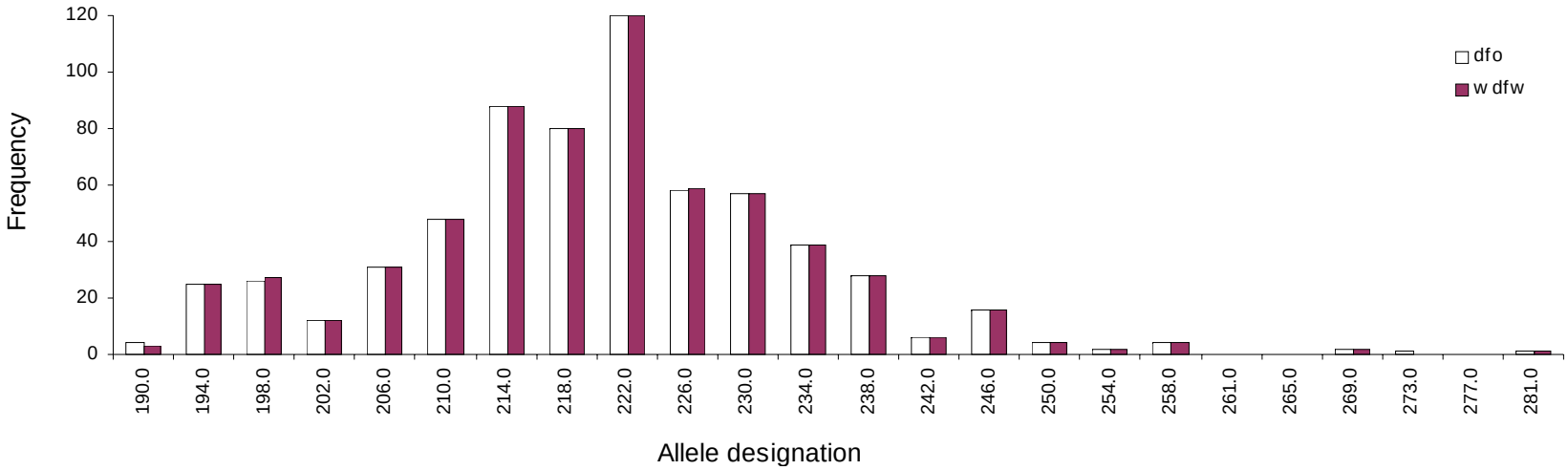

Figure 26: Concordance of designated alleles estimated by DFO and WDFW laboratories at Omy1011 in standardized and blind sample sets.

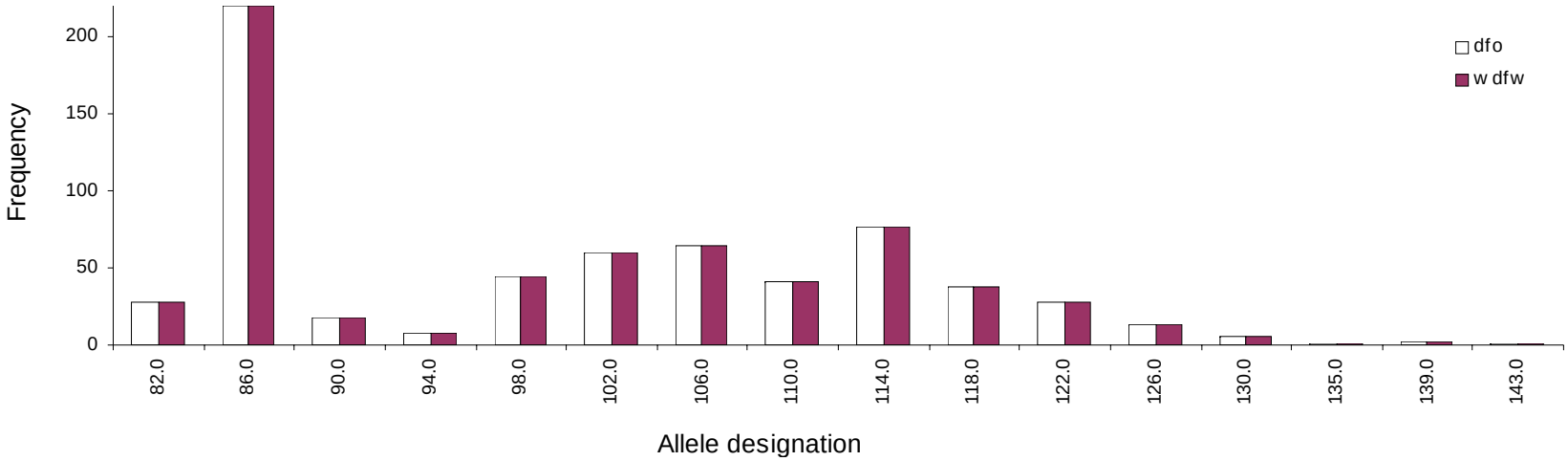

Figure 27: Concordance of designated alleles estimated by DFO and WDFW laboratories at Oki100 in standardized and blind sample sets.

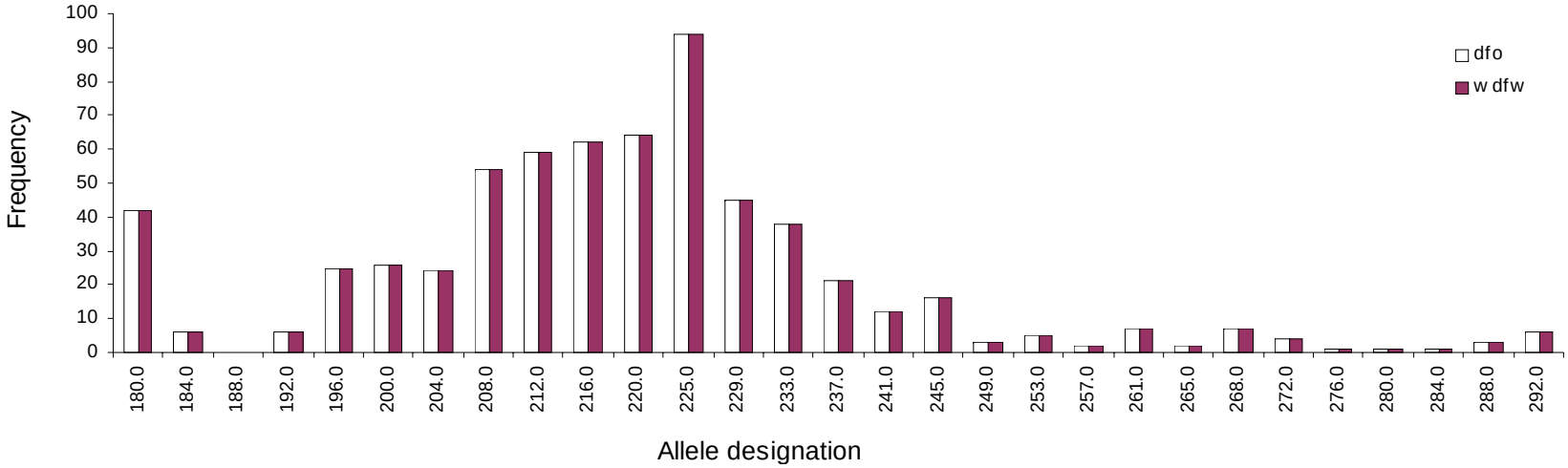

Figure 28: Concordance of designated alleles estimated by DFO and WDFW laboratories at One114 in standardized and blind sample sets.

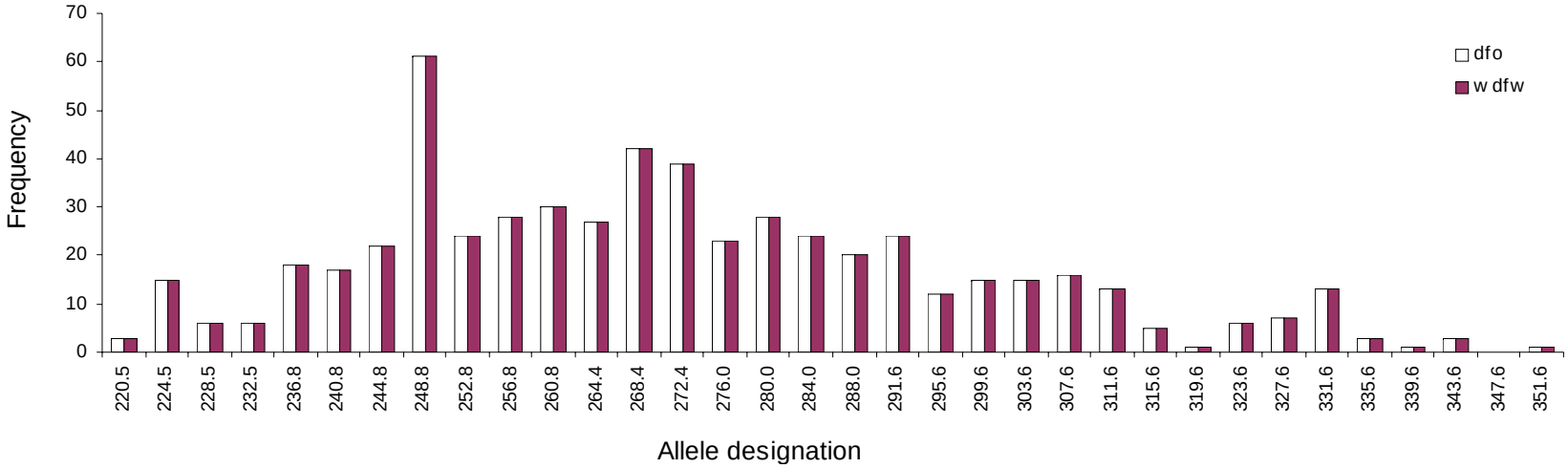

Figure 29: Concordance of designated alleles estimated by DFO and WDFW laboratories at Omm1070 in standardized and blind sample sets.

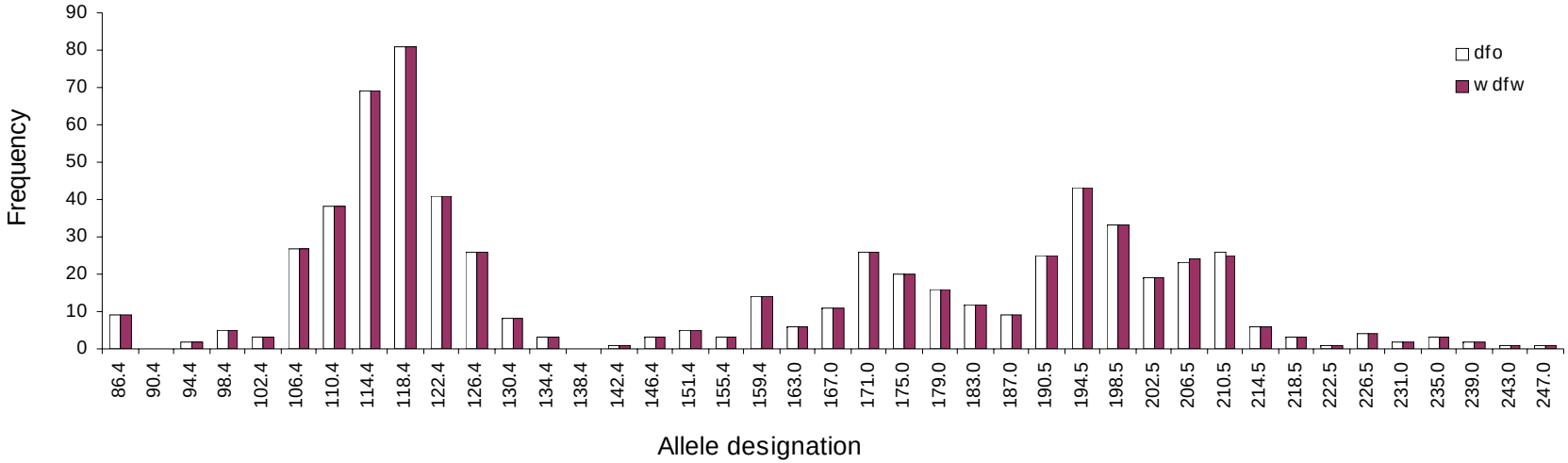

Figure 30: Concordance of designated alleles estimated by DFO and WDFW laboratories at Ots103 in standardized and blind sample sets.

Appendix A. BUDGET/EXPENDITURES SUMMARY: Joint development of a standardized DNA baseline for chum salmon in Washington and British Columbia – WDFW YR1 standardization of loci and alleles between labs (A-003)

Category	**Proposed**	**Actual**
Labor - Wages & Salaries	$15,950	$15,250
Labor - Employer Costs (benefits)	$3,509	$4,797
SUBTOTAL – LABOR	**$19,459**	**$20,047**
Site / Project costs - Site Supplies & Materials Consumable supplies and reagents	$2,743	$2,298
Travel Participate in standardization workshop in Nanaimo	$1,000	$857
Overhead Other overhead costs (WDFW indirect @ 29.3%)	$6,798	$6,798
TOTAL	**$30,000**	**$30,000**