Supplementary Methods

Parameter optimization process for finding ROH in *PLINK*:

All settings were chosen based on the premise of getting comparable output from the different datasets, as has been previously performed for human data (Ceballos et al., 2018). The resulting parameters achieved 93%-98% genome coverage across all datasets employed (apart from the very low-density RAD-seq data at 54%) for *Yellow*, and similarly for the 10 hihi.

Function: --homozyg-density
Here, the minimum density to consider a ROH (a ROH must have at least one SNP per this amount of kb) was set to 100kb for SNP chip data (Purfield et al., 2012; Purfield et al., 2017; Stoffel et al., 2021) and to 50kb for RAD-seq and the higher density data (Ceballos et al., 2018). However, for the larger datasets, varying the minimum density requirement had no impact on the ROH output.

Function: --homozyg-gap
Here, the maximum gap allowed between two SNPs (in order for them to be considered adjacent) was set to be 200kb, and 250kb for the RAD-seq dataset in order to account for the low SNP density, similar to other studies employing medium-density arrays (Purfield et al., 2017; Stoffel et al., 2021). We followed advice from Meyermans (et al., 2020) to minimize gap length while maintaining maximal genome coverage. For the larger datasets, varying the gap size had no impact on the ROH output.

Function: --homozyg-kb
Here, we chose 300 as the minimum length in kb that a run must have to be called as a ROH as this setting has been chosen for a comparable study using human data (Ceballos et al., 2018) as well as in ROH studies of wild populations (Hooper et al., 2020). We report the total length of all ROH found in Table S3 and S5 and total inbreeding F_{ROH} (using all ROH) for as well as $F_{ROH(>500Kb)}$ for *Yellow* (Table 1). However, only segments larger than 500Kb (e.g. Purfield et al., 2012) are used when calculating F_{ROH} for the ten hihi and in the remaining correlation analyses (Supplementary Table S10; comparison to *RZooRoH* in Supplementary Table Sheet S9).

Function: --homozyg-snp
Only ROH with at least this amount of SNPs are noted and the parameter values varied between datasets (between 14 and 100; e.g. Stoffel et al., 2021). A suitable threshold was estimated based on the number of SNPs expected in a 300kb window depending on the SNP density of the respective dataset, meaning that these values are similar to --homozyg-window-snp parameters.

Function: --homozyg-het*
As many other studies (Purfield et al., 2017; Ceballos et al., 2018; Hooper et al., 2020; Niskanen et al., 2020), we used the default *PLINK* setting of 5 heterozygous SNPs allowed in a final ROH, as the –homozyg-window-het filter already reduces the amount of heterozygous SNPs that the final ROH could possibly have. The default value seemed sufficient, especially for the high-density datasets, as an extra allowed heterozygous SNP (or more) did not influence the number of detected ROH or total ROH length by a lot. We checked this by allowing for 0 to 10 heterozygous sites per ROH while keeping the –homozyg-window-het parameter constant (see table below).

Function: --homozyg-window-snp
Here, the size of the sliding window in SNPs was calculated based on the number of SNPs expected in a 300kb window depending on the SNP density of the respective dataset, both for *Yellow* and the ten hihi, as a low minimal density setting (in kb/SNP) can lead to an incomplete genome coverage (Meyermans et al. 2020). Other studies report these numbers to be set between 30 and 100 for medium-size datasets (Purfield et al., 2017; Hooper et al., 2020; Niskanen et al., 2020)

Function: --homozyg-window-miss
Across all datasets, only one missing call was allowed in a ROH with *PLINK* as recommended in Meyermans et al., 2020 (which is a more conservative approach than most) to account for possible genotyping failure.

Function: --homozyg-window-het*
Here, we chose 1 or 2 as the number of heterozygous SNPs allowed in a ROH, depending on the size of the dataset and the quality of the calls. For the high-density low-coverage data, the benefit of tolerating heterozygous SNPs in order to control for genotyping calling errors outweighs the risk that two different homozygous segments are accidently merged.

Function: --homozyg-window-threshold
Here, we chose 0.05 as the minimum hit rate of all scanning windows containing a SNP to be eligible for inclusion in a ROH, as seen in many other inbreeding studies (e.g. Ferenčaković et al., 2013; Ceballos et al., 2018; Stoffel et al., 2021) and reviews (Meyermans et al., 2020).

*Example data for effect of changing –het or –window-het parameter:

SNP chip dataset						
- het	**window-het**	**#of ROH found**	**Length of ROH**	**window-het**	**#of ROH found**	**Length of ROH**
default	0	1256	133182	6	1460	609198
default	1	1357	180502	7	1695	712007
default	2	1321	224600	8	2045	783294
default	3	1261	303923	9	2537	847563
default	4	1266	390042	10	3167	883846
default	5	1343	491749			

SNP chip dataset						
window-het	**-het**	**#of ROH found**	**Length of ROH**	**-het**	**#of ROH found**	**Length of ROH**
1	0	1256	133182	6	1356	180466
1	1	1268	146648	7	1357	180502
1	2	1291	175610	8	1357	180502
1	3	1337	179164	9	1357	180502
1	4	1351	179745	10	1357	180502
1	5	1352	179890			

Supplementary Material

Table S1: Per-individual mapped reads, absolute and relative missingness for the RAD-seq data (26,447 SNPs) and the low-coverage whole-genome data (2,018,863 SNPs) of the ten sampled hihi from Tiritiri Matangi and Te Hauturu-o-Toi. Values in bold represent high rates of missingness or especially low per-site coverage relative to the other hihi.

	RAD-seq data			lcWGS data		
ID	**Total number of raw reads**	**Number of sites missing**	**frequency missingness**	**Total number of raw reads**	**Number of sites missing**	**frequency missingness**
Hihi_01	5,411,909	4,211	0.159	**45,115,882**	**160,625**	**0.080**
Hihi_02	5,661,034	4,604	0.174	**63,790,336**	**64,534**	**0.032**
Hihi_03	11,488,959	1,956	0.074	70,141,628	47,616	0.024
Hihi_04	12,127,620	1,582	0.060	88,507,344	38,708	0.019
Hihi_05	**3,691,785**	**9,439**	**0.357**	100,136,630	35,399	0.018
Hihi_06	9,384,072	2,491	0.094	86,780,654	19,832	0.010
Hihi_07	11,002,286	1,655	0.063	91,194,314	35,660	0.018
Hihi_08	**2,647,254**	**14,286**	**0.540**	90,509,738	18,899	0.009
Hihi_09	9,726,802	2,417	0.091	117,320,750	18,435	0.009
Hihi_10	7,809,480	4,660	0.176	126,397,278	12,566	0.006

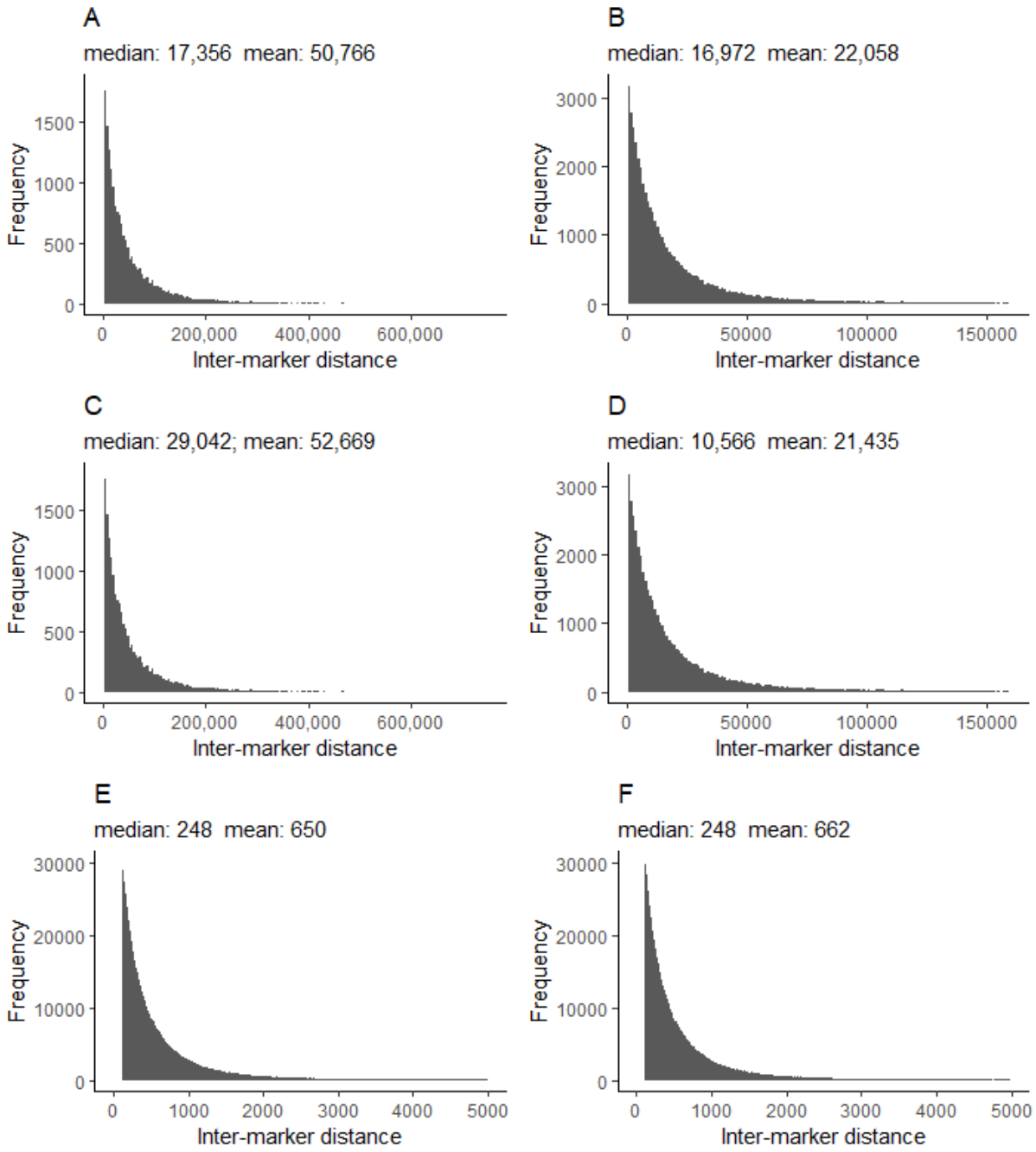

Figure S1: Distribution of inter-marker distances for four datasets of the assembly male hihi *Yellow* generated using the package *dplyr* in R. From top left to bottom right: A: RAD-seq, B: SNP array, C: WGS, D: Combined dataset. Grey line: median. Red line: mean. As expected with low-density data, the average inter-marker distance for the RAD-seq dataset is much higher than in the lcWGS data. Note the different scaling of the x-axis per plot.

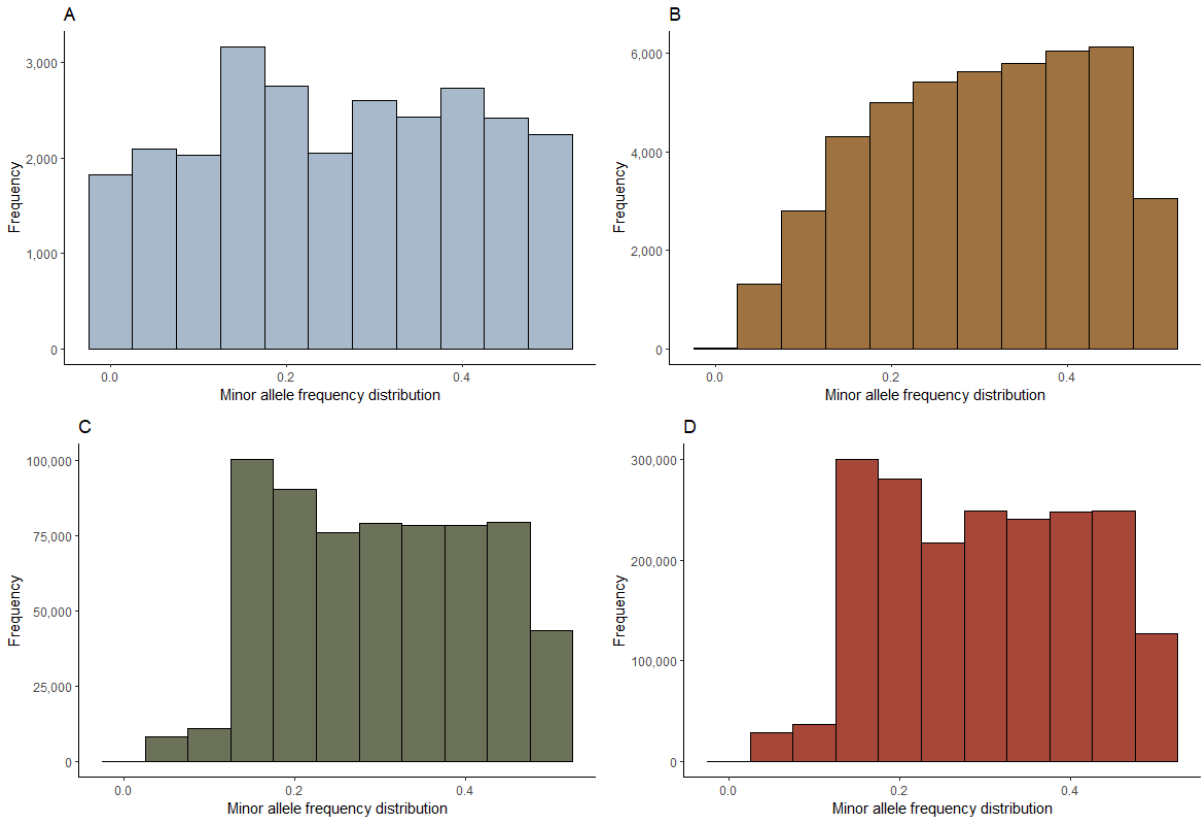

Figure S2: Minor allele frequency spectra for the (A) RAD-seq, (B) SNP array, (C) lcWGS8 and (D) lcWGS datasets for the ten hihi. Note that the two lcWGS datasets were filtered to have a minimum minor allele frequency of 0.02.

Table S2: Overview of PLINK's ROH analysis settings employed across all SNP datasets. We required a minimum length of 300 kb for a reported ROH in order to exclude short ROH deriving from population linkage disequilibrium, and allowed for one missing site per ROH (Hillestad et al., 2017; Meyermans et al., 2020). The custom settings for the other parameters were adjusted for the SNP density of a dataset, as detailed in Supplementary Tables S3, S4 and S5.

Parameters for *PLINK*s --homozyg function	**Name of function**	**Value**
Minimum density to consider a ROH (a ROH must have at least one SNP per this amount of kb)	--homozyg-density	custom
Maximum gap allowed between two SNPs (in order for them to be considered adjacent)	--homozyg-gap	custom
Minimum length in kb that a run must have to be called as a ROH	--homozyg-kb	300
Only ROH with at least this amount of SNPs are noted	--homozyg-snp	custom
Amount of heterozygous calls a final ROH can have	--homozyg-het	*default*
Size of the sliding window in SNPs	--homozyg-window-snp	custom
Number of missing calls allowed in a scanning window hit	--homozyg-window-miss	1
Number of heterozygous SNPs allowed in a scanning window hit	--homozyg-window-het	custom
Minimum hit rate of all scanning windows containing a SNP to be eligible for inclusion in a ROH	--homozyg-window-threshold	0.05

Table S3: Overview of ROH analysis input parameters as well as output summary for the assembly male Yellow. *Settings not mentioned here are the same as in Table S2. At the bottom, six example contigs with their respective detected ROH per dataset are listed.*

Dataset	Full genome	Combined	WGS	SNP array	RAD-seq
Genotyped sites	904,228,112	1,593,073	1,562,384	46,136	18,415
Average density: 1 site/x bases	1	568	580	19,600	49,100
#sites expected per 300kb	300,000	525	517	15	6
***PLINK* ROH settings**					
--homozyg-density	50	50	50	100	50
--homozyg-gap	200	200	200	200	250
--homozyg-snp	300,000	100	100	30	10
--homozyg-window-het	2	2	2	1	1
--homozyg-window-snp	300,000	50	50	15	6
Total #ROH found	270	290	300	223	285
Total length (kb)	201,000	199,335	207,481	236,390	186,613
$F_{(ROH)}$ (≥500kb)	0.1487	0.1378	0.1437	0.2240	0.1299
#sites in ROH	176,447,917	263,068	290,384	11,421	6,021
Length of ROH for six examples contigs (kb) (numbers in brackets indicate the number of ROHs detected, if different from 1)					
Contig_275	6067	5929	6027	6055	3666
Contig_527	1725 (3)	1737 (3)	1734 (3)	1490 (2)	512 (1)
Contig_863	570	574	588	-	421
Contig_741	2229	1943	2232	2835	2064
Contig_433	1594	1621	1623	1738	1368
Contig_1404	486	504	504	-	553

Table S4: In Excel spreadsheet: Overview of ROH analysis input parameters as well as output summary for down-sampled datasets for the assembly male Yellow. *Settings not mentioned here are the same as in Table S2.*

Table S5: Overview of ROH analysis for the ten hihi in PLINK. Included were all datasets of the low-coverage whole-genome sequencing, the SNP array data and the RAD-seq data. PLINK parameter settings were adjusted based on the density of the dataset. Settings not mentioned here are the same as in Table S2. Reported are the total number of ROH found and the total length of ROH per individual 01-10 with each of the datasets.

Dataset		**WGS**	**WGS6**	**WGS7**	**WGS8**	**WGS9**	**ARRAY**	**RAD**
Sites genotyped		1,974,836	1,746,437	1,302,507	644,631	195,184	45,553	26,447
Sites genotyped in all hihi		1,658,958	1,507,697	1,159,979	593,077	184,539	45,553	26,447
1 site/x bases		545	600	780	1,525	4,999	19,850	34,191
--homozyg-density		50	50	50	50	50	100	50
--homozyg-gap		200	200	200	200	200	200	250
--homozyg-snp		100	100	100	90	50	30	14
--homozyg-window-het		2	2	2	2	2	1	1
--homozyg-window-snp		50	50	50	45	25	15	8
Hihi_01	Total #ROH:	**47**	**60**	88	186	281	155	283
	Tot. length (kb):	**22,749**	**29,136**	42,596	96,141	180,740	176,157	176,206
Hihi_02	Total #ROH:	**168**	**201**	237	266	288	152	262
	Tot. length (kb):	**81,375**	**104,432**	127,660	152,112	178,339	188,171	165,237
Hihi_03	Total #ROH:	227	260	283	274	247	173	297
	Tot. length (kb):	116,567	145,353	169,498	182,371	185,542	222,029	190,851
Hihi_04	Total #ROH:	282	310	294	273	263	140	292
	Tot. length (kb):	153,264	186,190	194,303	206,371	212,330	213,520	199,885
Hihi_05	Total #ROH:	356	366	342	308	304	168	**129**
	Tot. length (kb):	193,488	229,839	236,223	243,018	246,398	258,402	**73,773**
Hihi_06	Total #ROH:	311	303	296	273	244	158	283
	Tot. length (kb):	179,780	183,574	187,584	182,620	173,564	172,024	182,908
Hihi_07	Total #ROH:	298	293	275	270	274	133	290
	Tot. length (kb):	175,184	186,171	182,308	183,430	185,238	180,502	187,588
Hihi_08	Total #ROH:	344	320	295	282	251	155	**18**
	Tot. length (kb):	223,218	222,622	220,088	220,480	207,454	203,972	**8,206**
Hihi_09	Total #ROH:	374	344	317	291	286	145	313
	Tot. length (kb):	255,982	267,331	256,167	252,392	255,175	267,979	212,865

Hihi_10	Total #ROH:	319	283	271	274	265	164	256
	Tot. length (kb):	197,348	189,162	178,892	180,959	181,257	203,735	147,921

Table S6: In Excel spreadsheet. Correlations between missingness and the number and total length of ROH detected for the ten hihi.

Table S7. In Excel spreadsheet. Contig sizes and the number and length of ROHs detected for of the ten hihi.

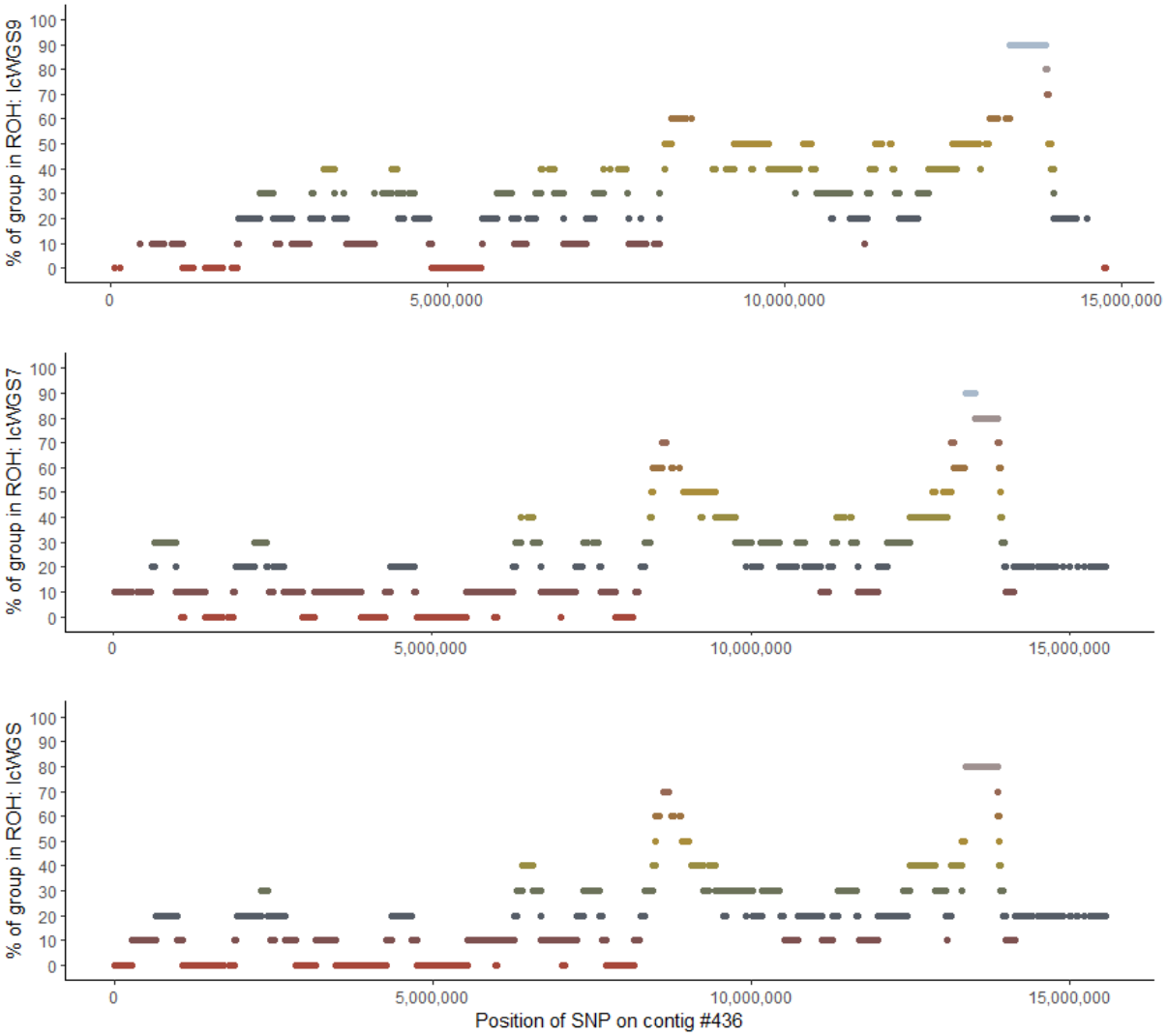

Figure S3: Location of runs of homozygosity (ROH) for one example contig (#436) that had ROH detected across all seven datasets, with ROH landscapes shown for the lcWGS9, lcWGS7 and lcWGS datasets. Displayed are the percentage of the ten birds (y-axis) that have this SNP (x-axis) involved in a run of homozygosity. The more individuals share the ROH, the higher the SNP is located in the plot. Red dots at the bottom of the plot mean that those SNPs were not involved in a ROH in any bird. Other datasets (RAD-seq, SNP array and lcWGS8) are displayed in Figure 3.

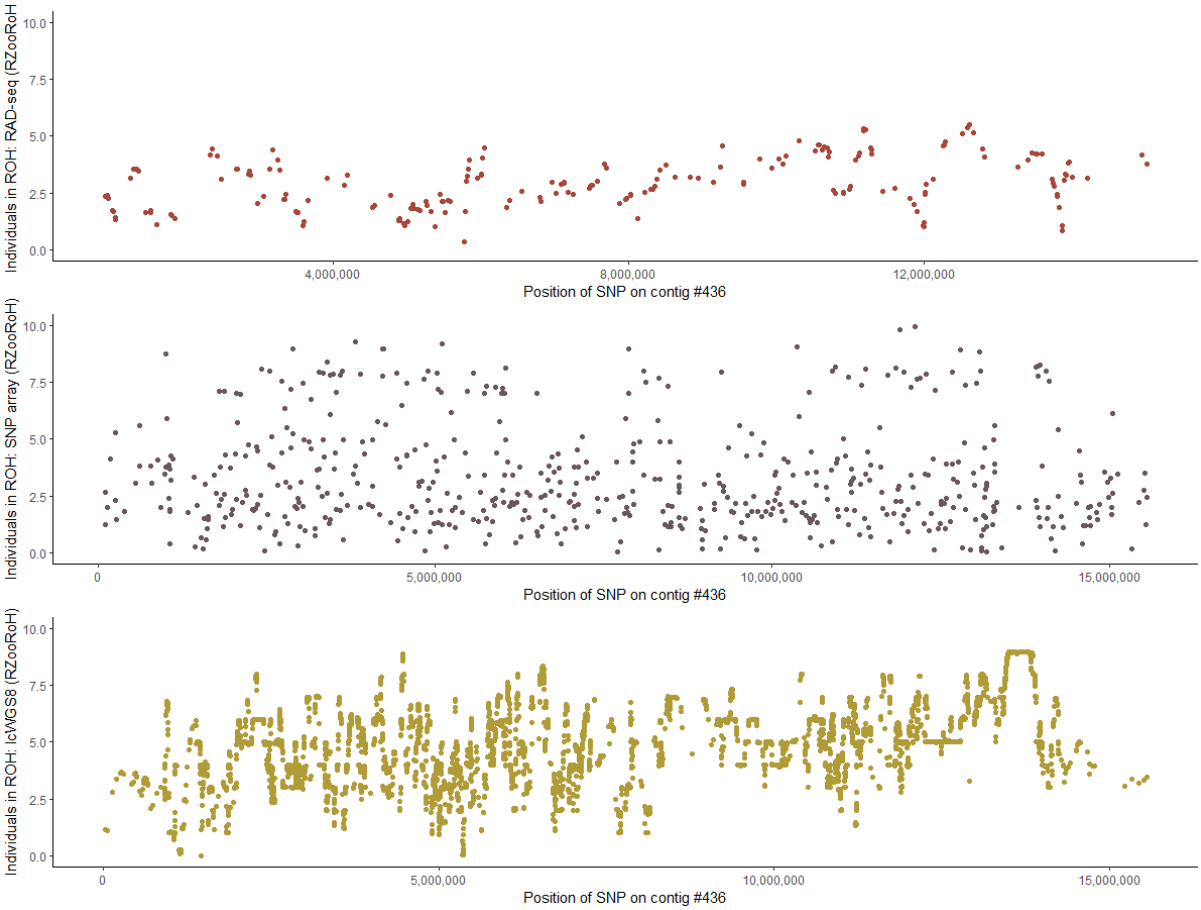

Figure S4: Displayed is the RZooRoH HBD landscape for the same contig (#436) for the ten hihi, showing the results from the RAD-seq, SNP array and WGS8 datasets (as in Figure 3 in the main document). The y-axis labelling "Individuals in ROH" represents the summed probability of the SNP to be in a HBD segment across individuals.

Table S8: In Excel spreadsheet. Inbreeding measures F_{IBC}, F_{HET} and F_{ROH} for the ten hihi.

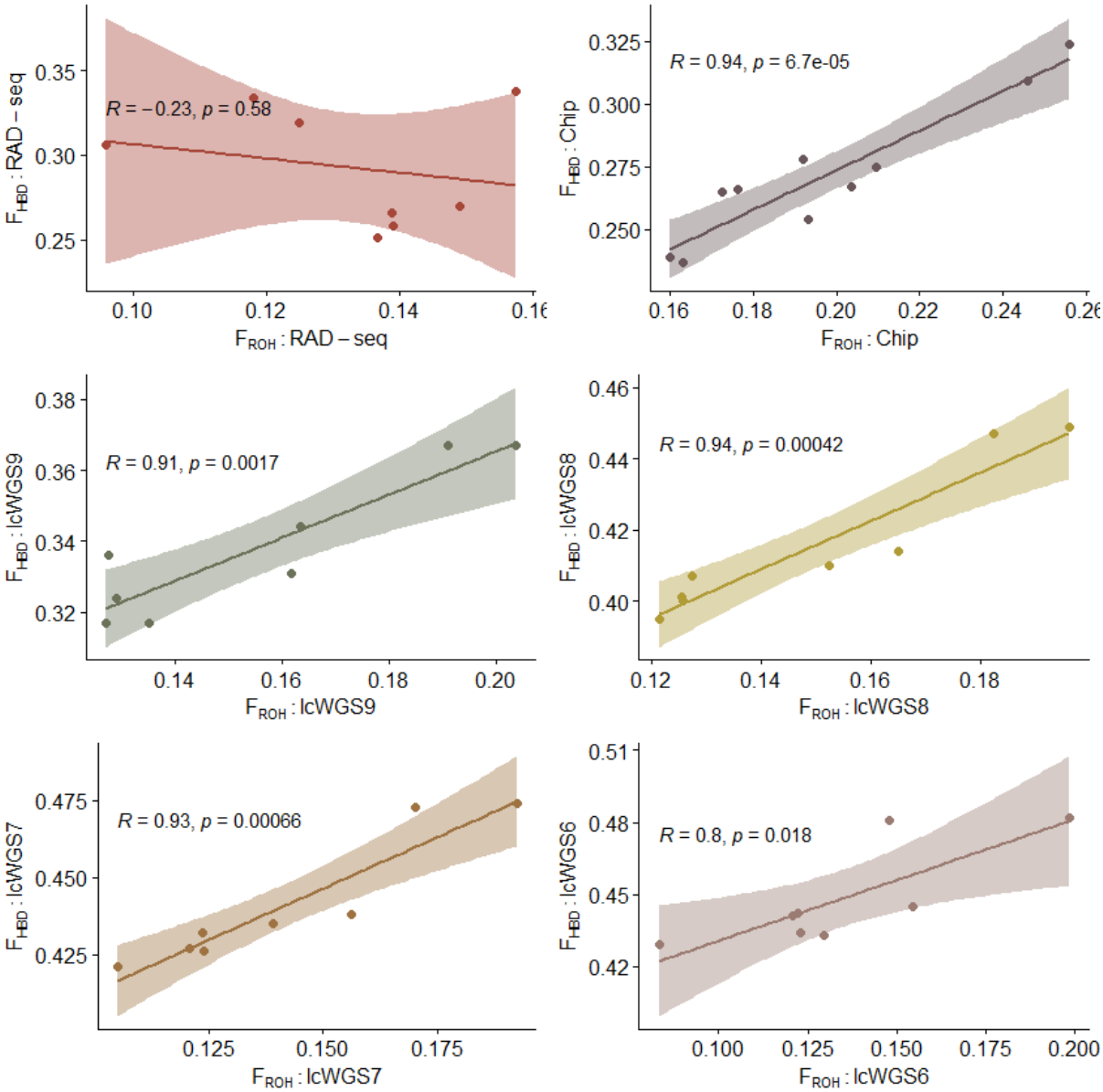

Figure S3: Correlation of inbreeding values based on PLINK (excluding ROH segments <500Kb) and RZooRoH HBD classes output. Two individuals were excluded in the RAD-seq plot for their high levels of genotype missingness (>30%) and overall two individuals are missing for the two larger lcWGS datasets due to low coverage. A list of inbreeding values can be found in the Supplementary Excel Sheet S9.

Table S9: In Excel spreadsheet. Correlations of F_{ROH} (PLINK) with $F_{RZooRoH}$ for the ten hihi.

Table S10: Overview of the inbreeding measures and multi-locus-heterozygosity for all ten hihi individuals when using the whole-genome low coverage dataset with a vcftools –min-meanDP 8 filter employed (i.e. lcWGS8). In InbreedR, the function mlh *estimated MLH as the total number of heterozygous loci in an individual divided by the number of loci typed in the focal individual. The function* sMLH *calculated MLH as the total number of heterozygous loci in an individual divided by the sum of average observed heterozygosities in the population over the subset of loci successfully typed in the focal individual.* F_{ROH} *was calculated by dividing the total ROH length (segments larger than 500kb) by the assembly genome size of 912Mb. F measures for the other datasets are available in Supplementary Table S8.*

ID	F_{ROH}	$F_{RZooRoH}$	F_{HET}	F_{IBC}	MLH
Hihi_01	0.050	0.429	0.292	0.250	0.370
Hihi_02	0.092	0.391	0.081	0.090	0.425
Hihi_03	0.121	0.395	0.037	0.051	0.437
Hihi_04	0.152	0.410	0.048	0.034	0.437
Hihi_05	0.182	0.447	0.100	0.089	0.418
Hihi_06	0.127	0.407	-0.004	0.003	0.462
Hihi_07	0.126	0.400	-0.008	0.017	0.456
Hihi_08	0.165	0.414	0.012	-0.013	0.463
Hihi_09	0.196	0.449	0.040	0.046	0.438
Hihi_10	0.125	0.401	-0.053	-0.027	0.474

Table S11: In Excel spreadsheet. Correlations of F_{IBC}, F_{HET}, $F_{RZooRoH}$ and F_{ROH} with MLH for the ten hihi.

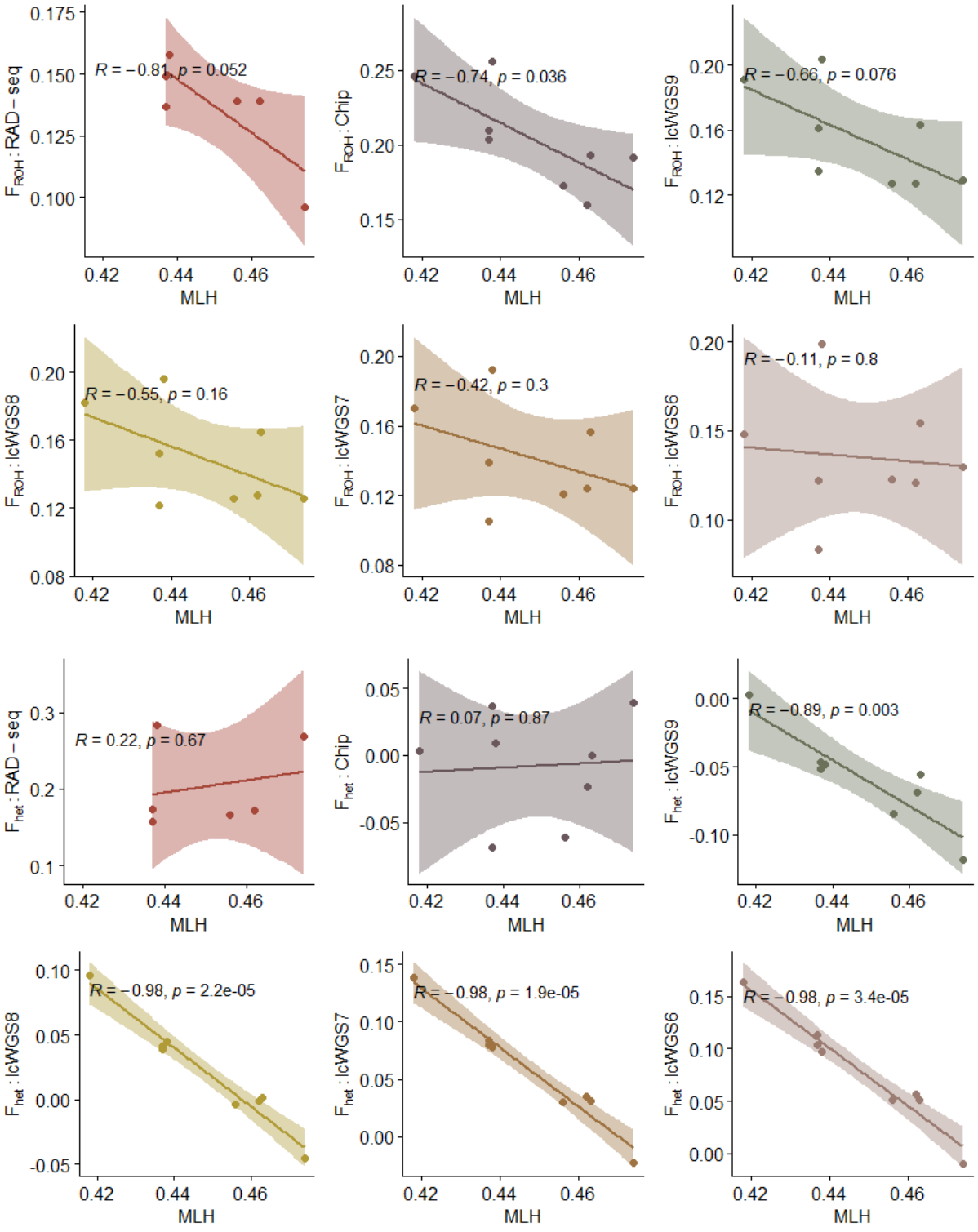

Figure S6: Displayed are the Pearson correlations between per-individual WGS8-based multi-locus heterozygosity (based on InbreedR) and two different measures of F (F_{HET}, F_{ROH}) that were estimated in this paper (F_{IBC} is strongly correlated with F_{HET}, hence not plotted). Two individuals were excluded in the RAD-seq plot for their high levels of genotype missingness (>30%) and overall two individuals are missing for the two larger lcWGS datasets due to low coverage. Each colour represents one of the different datasets for the ten initial hihi, from low-density RAD-seq data to more than 1.5m markers in the largest dataset (WGS6). FROH is based on homozygous segments larger

than 500kb. All correlations and regression plotting was performed in R studio. A full list of MLH correlations can be found in Supplementary Table S11.

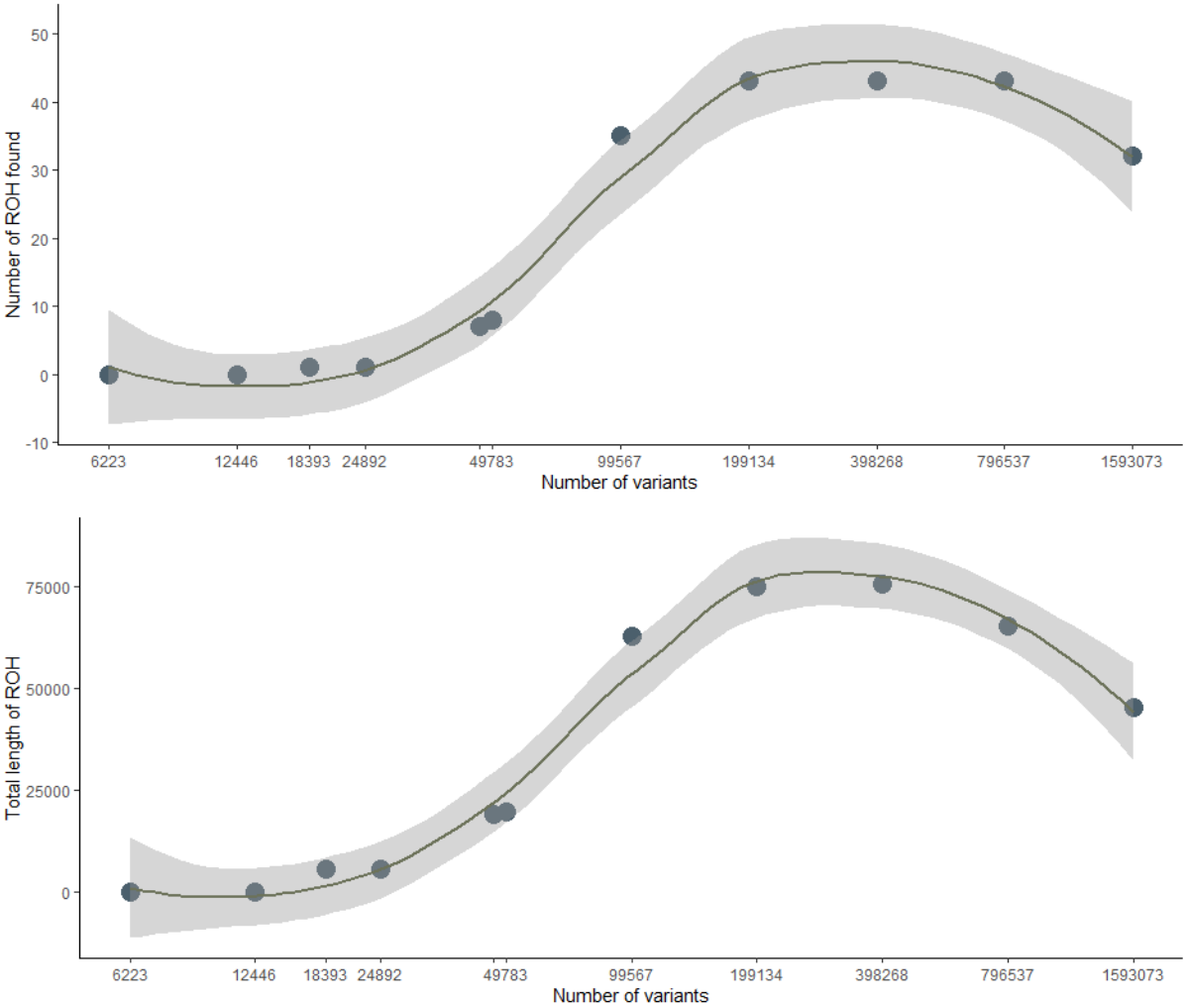

Figure S7: Scatterplot visualizing the ROH output for our down-sampled combined dataset for the assembly male when using default settings in PLINK to detect ROH. Datasets were randomly down-sampled to 1/2, 1/4,…, 1/256 markers, with sampling repeated ten times; the mean across datasets is plotted. The number of variants are plotted on a log scale. The grey shaded area represents the standard error as calculated with the geom_smooth(method = "loess") function in ggplot2 in R. The highest number of ROH were found in the marker-density range from 100,000 to 800,000 SNPs.

Ceballos, F. C., Hazelhurst, S., & Ramsay, M. (2018). Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics, 19*(1), 106. doi:10.1186/s12864-018-4489-0

Ferenčaković, M., Solkner, J., & Curik, I. (2013). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution, 45*(1), 42. doi:10.1186/1297-9686-45-42

Hillestad, B., Woolliams, J. A., Boison, S. A., Grove, H., Meuwissen, T., Vage, D. I., & Klemetsdal, G. (2017). Detection of runs of homozygosity in Norwegian Red: Density, criteria and genotyping quality control. *Acta Agriculturae Scandinavica Section A - Animal Science, 67*(3-4), 107-116. doi:10.1080/09064702.2018.1501088

Hooper, R., Excoffier, L., Forney, K. A., Gilbert, M. T. P., Martin, M. D., Morin, P. A., . . . Foote, A. D. (2020). Runs of homozygosity in killer whale genomes provide a global record of demographic histories. *bioRxiv*, 2020.2004.2008.031344. doi:10.1101/2020.04.08.031344

Meyermans, R., Gorssen, W., Buys, N., & Janssens, S. (2020). How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics, 21*(1), 94. doi:10.1186/s12864-020-6463-x

Niskanen, A. K., Billing, A. M., Holand, H., Hagen, I. J., Araya-Ajoy, Y. G., Husby, A., . . . Jensen, H. (2020). Consistent scaling of inbreeding depression in space and time in a house sparrow metapopulation. *Proceedings of the National Academy of Sciences, 117*(25), 14584-14592. doi:10.1073/pnas.1909599117

Purfield, D. C., Berry, D. P., McParland, S., & Bradley, D. G. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics, 13*(1), 70. doi:10.1186/1471-2156-13-70

Purfield, D. C., McParland, S., Wall, E., & Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One, 12*(5), e0176780. doi:10.1371/journal.pone.0176780

Stoffel, M. A., Johnston, S. E., Pilkington, J. G., & Pemberton, J. M. (2021). Mutation load decreases with haplotype age in wild Soay sheep. *Evolution Letters, n/a*(n/a). doi:https://doi.org/10.1002/evl3.229