The American Journal of Human Genetics, Volume *93*

Supplemental Data

Whole-Genome Sequencing Uncovers the Genetic Basis of Chronic Mountain Sickness in Andean Highlanders

Dan Zhou, Nitin Udpa, Roy Ronen, Tsering Stobdan, Junbin Liang, Otto Appenzeller, Huiwen W. Zhao, Yi Yin, Yuanping Du, Lixia Guo, Rui Cao, Yu Wang, Xin Jin, Chen Huang, Wenlong Jia, Dandan Cao, Guangwu Guo, Jorge L. Gamboa, Francisco Villafuerte, David Callacondo, Jin Xue, Siqi Liu, Kelly A. Frazer, Yingrui Li, Vineet Bafna, and Gabriel G. Haddad

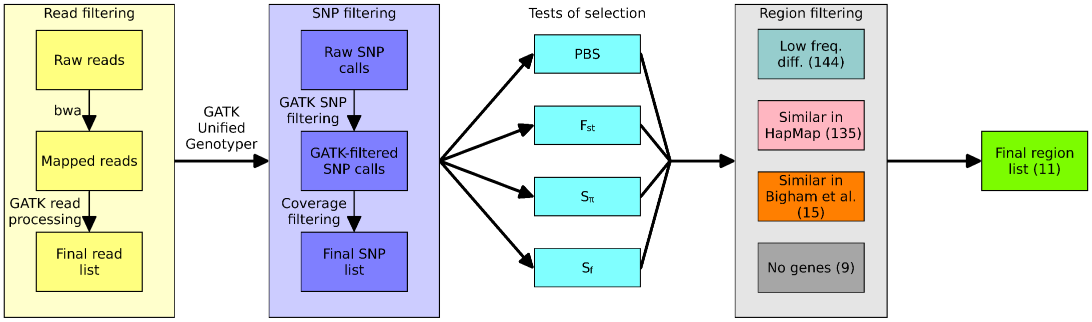


Figure S1. Computational analysis workflow. The raw sequence reads were mapped using BWA, followed by indel realignment, duplicate marking, and quality score recalibration using the GATK pipeline. Variants were then called and filtered using the GATK UnifiedGenotyper. After applying additional variant filters to account for the differences in coverage between our subjects and the control populations, we applied four complementary tests of selection to identify 314 regions as candidates for undergoing hypoxia-specific selective stress. Of these, we prioritized 11 regions for further experimental validation.

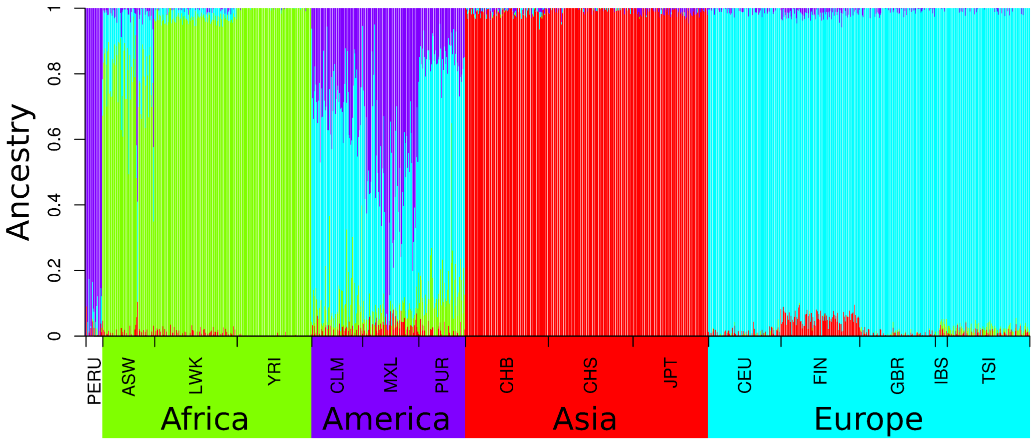


Figure S2. ADMIXTURE analysis (4 clusters) applied to the 1000 Genomes populations jointly with our Andean highlander subjects. The Andean highlanders (leftmost) show no signs of significant shared ancestry with Europeans or Africans, but do show varying degrees of shared ancestry with the American populations (CLM, MXL, PUR). As lowlander controls for our cross-population tests of selection, we used the closest population (MXL). As an out-group for the PBS test, we used a distant African population (LWK).

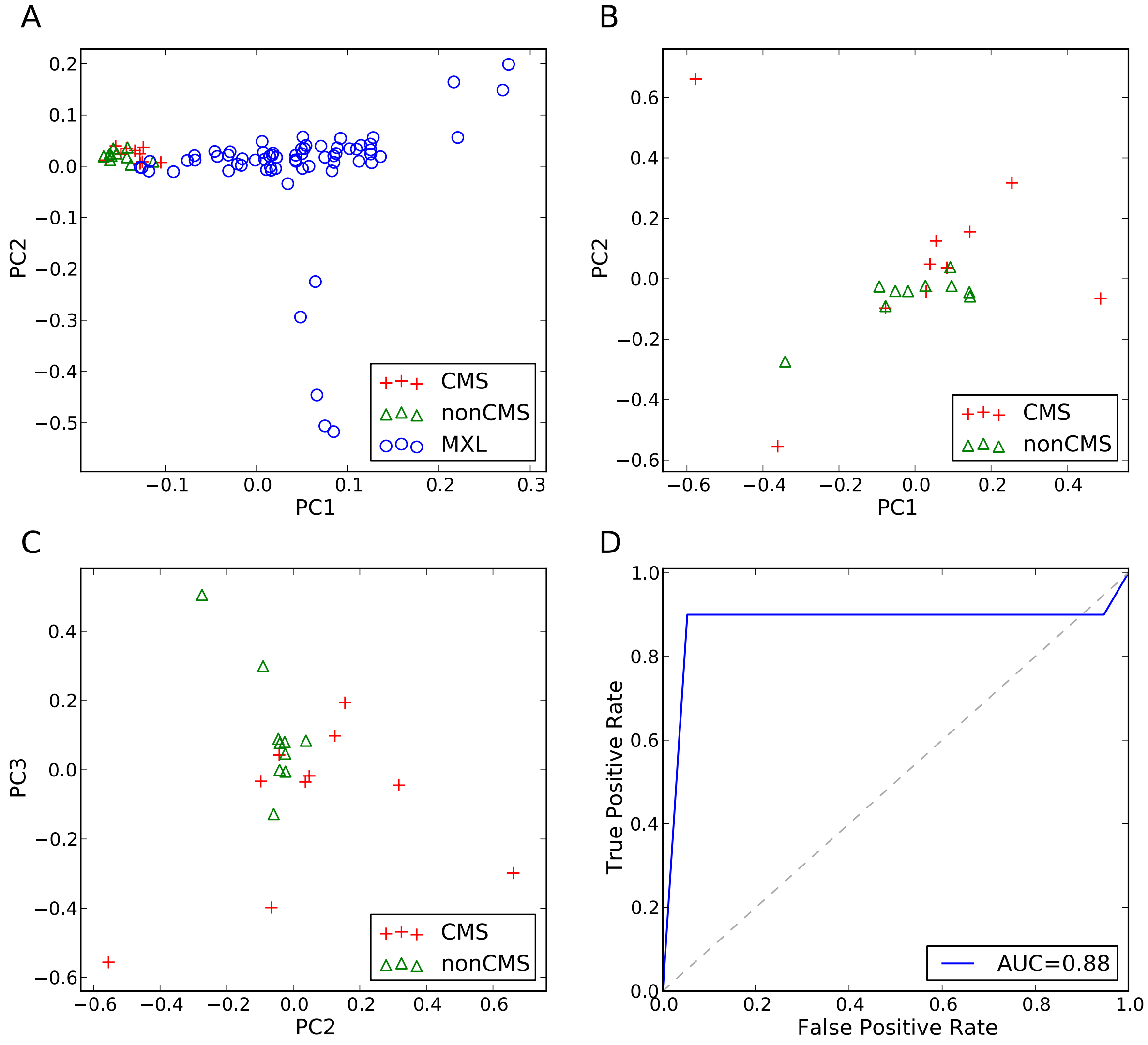


Figure S3. PCA plots of the highlander subjects and lowlander controls. (A) First and second principal components from PCA performed on the 20 Andean subjects, and MXL controls. While we observe clustering of the Andeans relative to MXL, there is no hyperplane separating the two populations, further illustrating the closeness (and thus, appropriateness) of the MXL control. (B) First and second principal components from PCA performed only on the 20 Andeans subjects. (C) Second and third principal components from PCA performed on the 20 Andeans subjects. The top three principal components account for a similar amount of the total variance. The third principal component provides a fair separation (P=0.04, Wilcoxon test) between the CMS and non-CMS subjects, indicating a certain degree of population substructure between the two phenotypic groups. (D) ROC of a linear model trained on the first, second, and third principal components of the PCA from (B,C). The observed AUC of 0.88 has P-value of 0.033 when permuting class labels (1000 permutations).

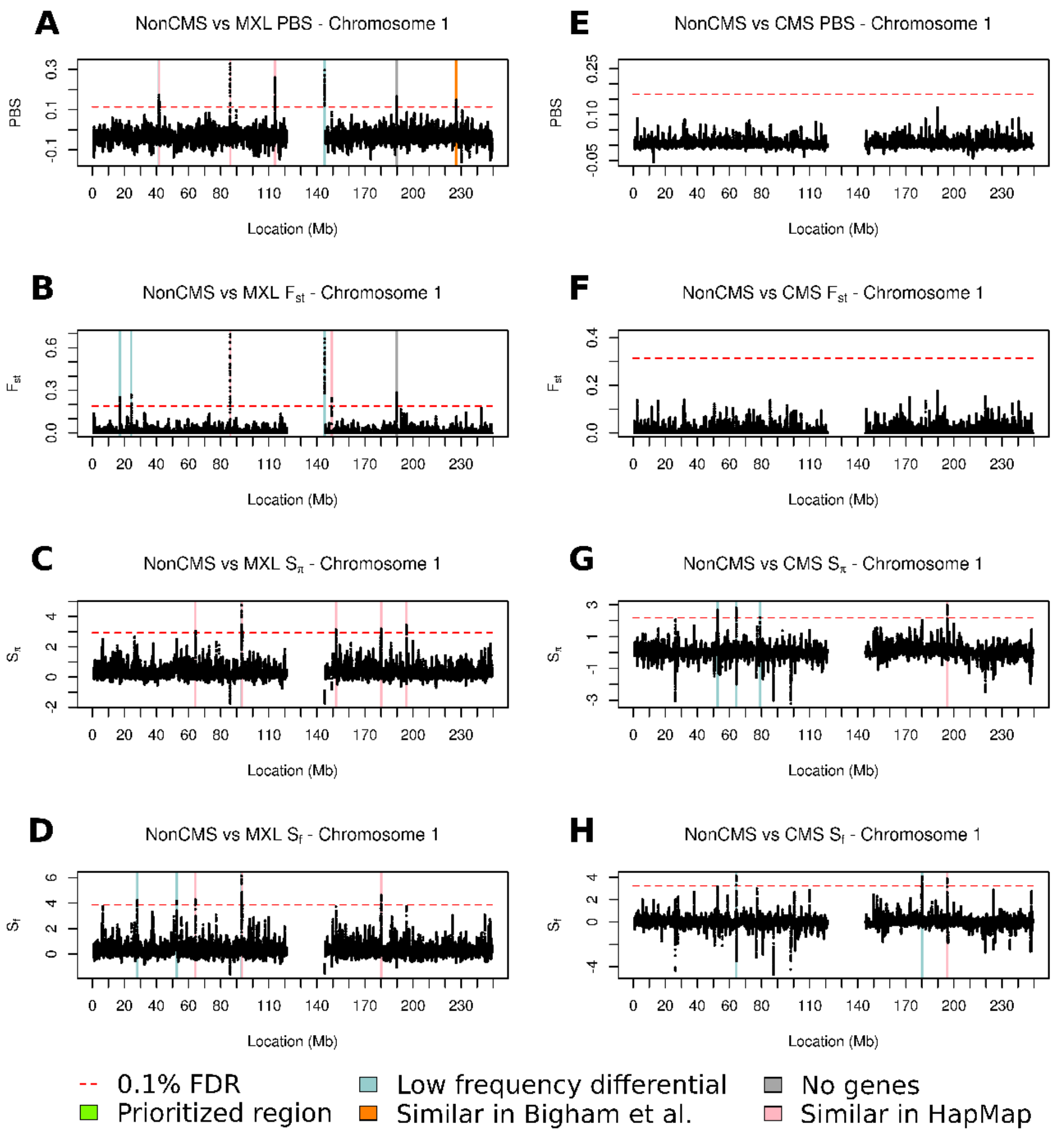


Figure S4. Test statistic values on chromosome 1, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

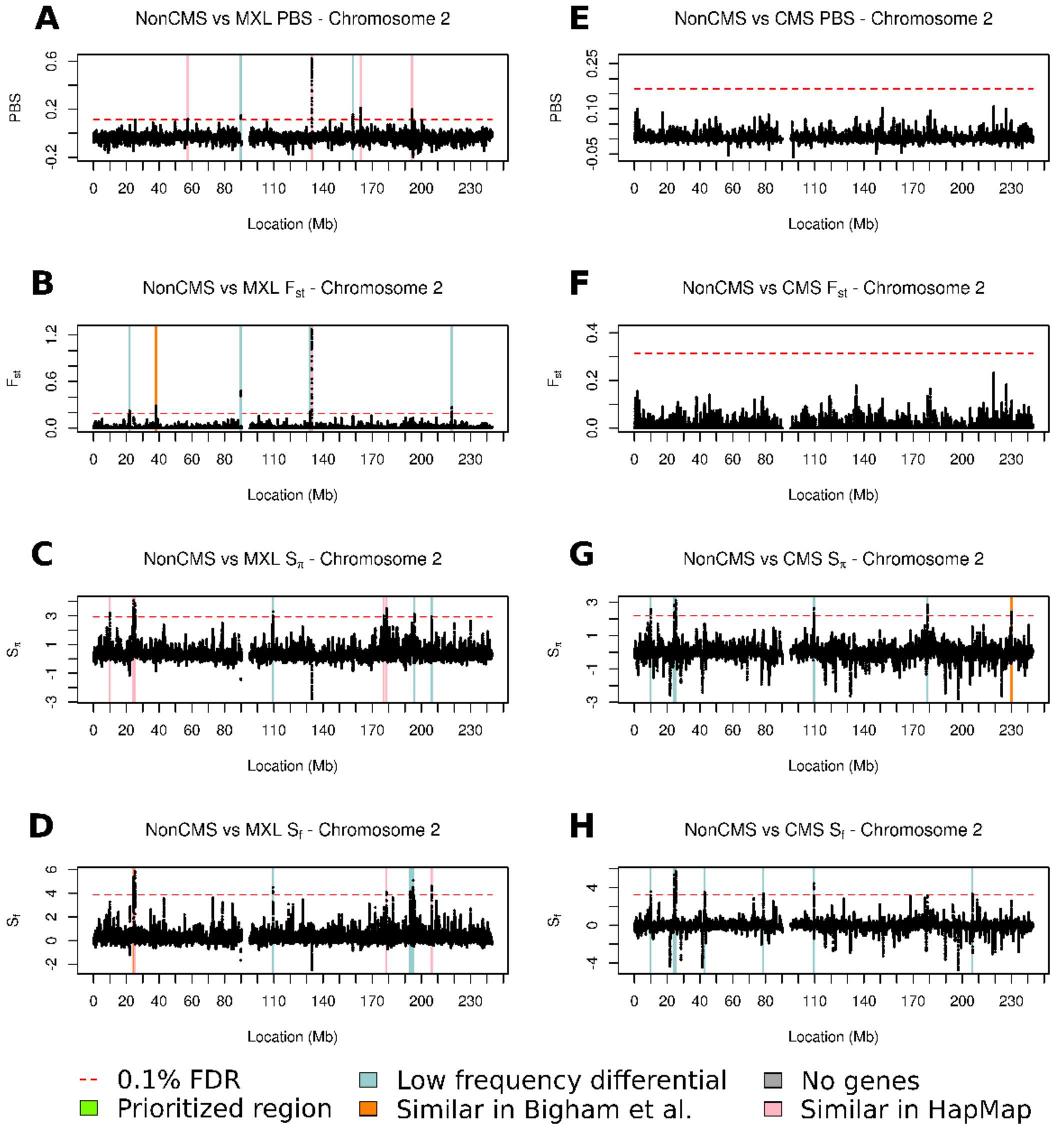


Figure S5. Test statistic values on chromosome 2, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

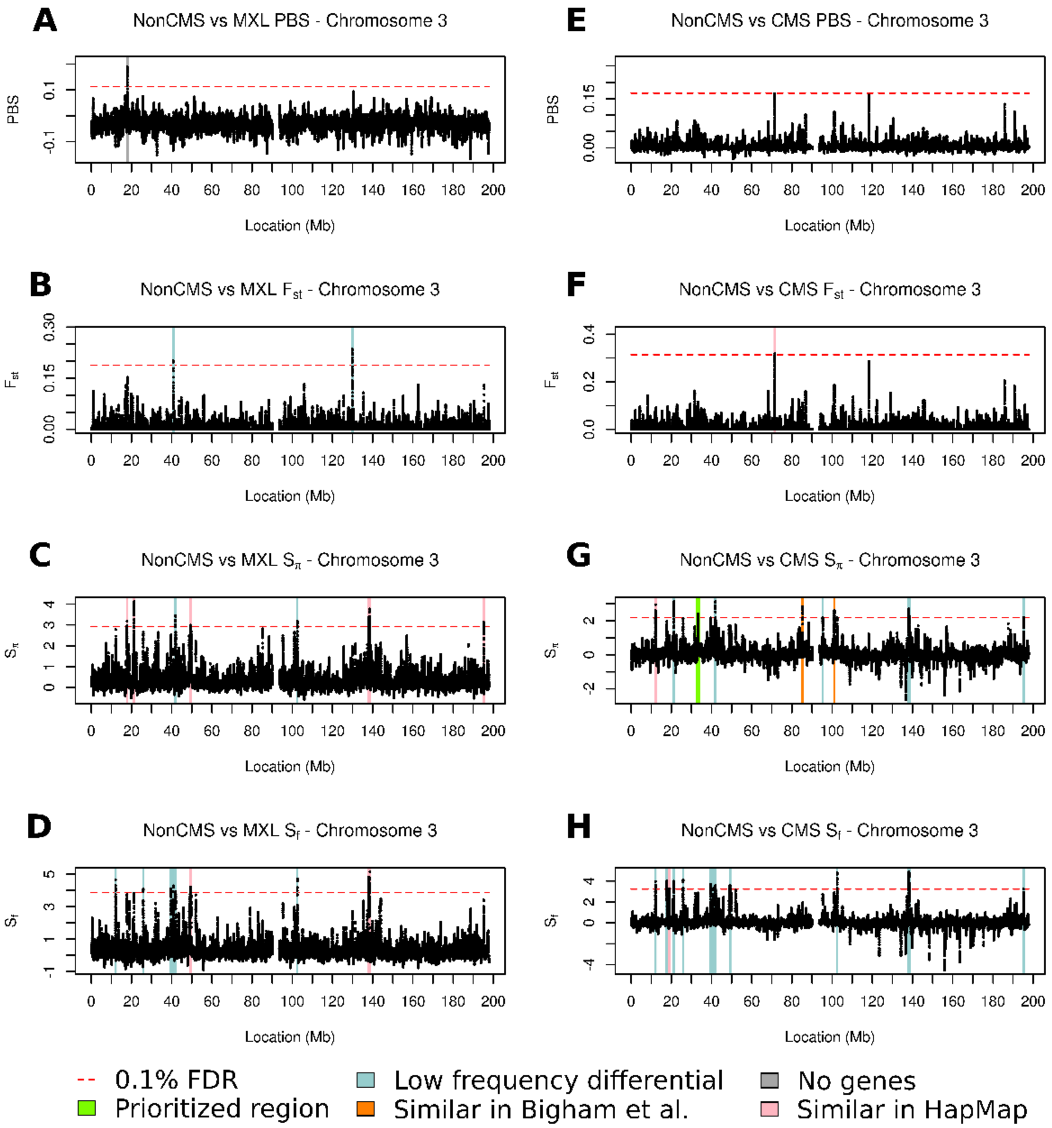


Figure S6. Test statistic values on chromosome 3, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. CMS S_π statistic extends from positions 33,254,596-33,314,596, contains 1 gene (*SUSD5*), and has maximum S_π statistic value of 2.195.

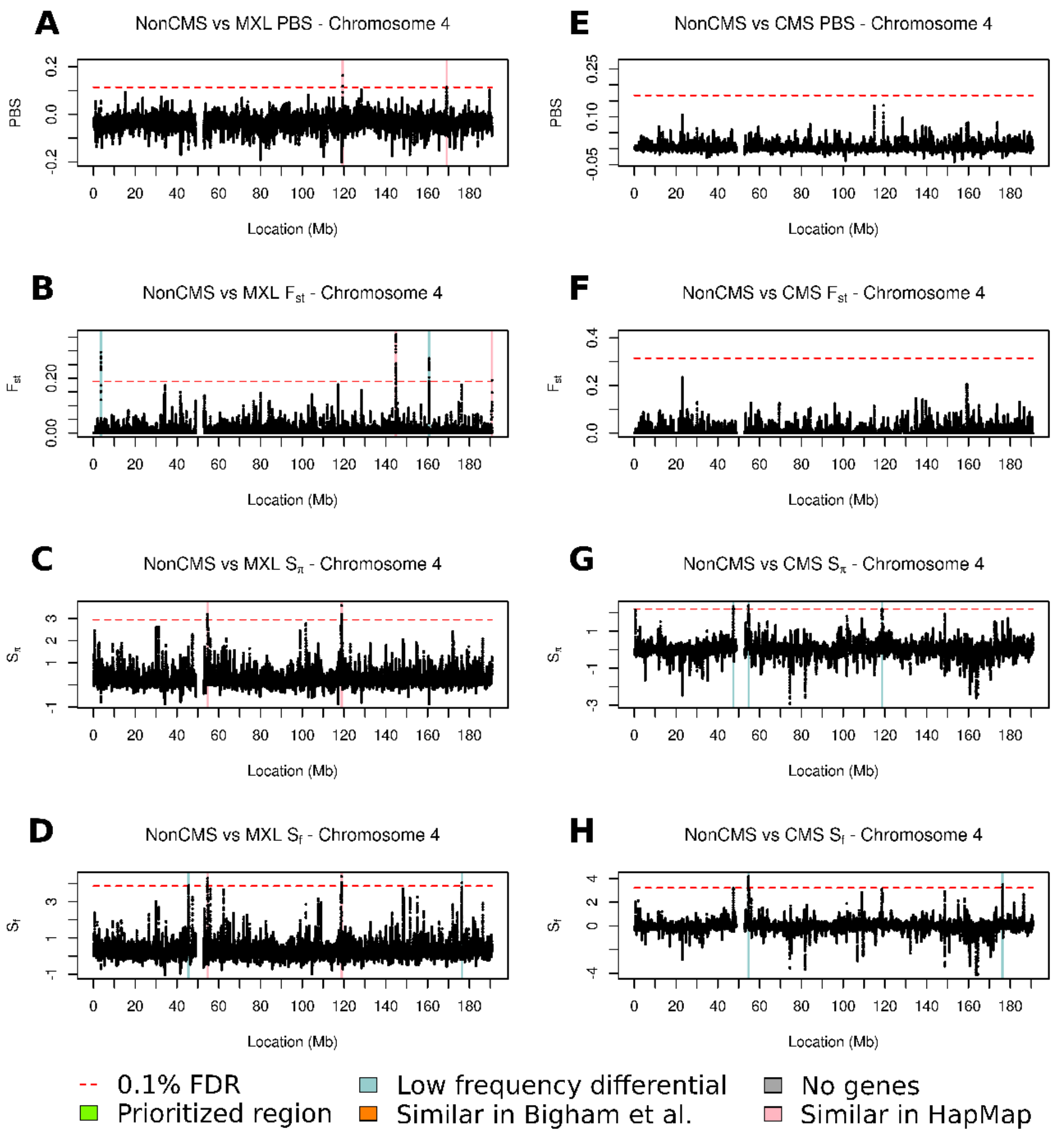


Figure S7. Test statistic values on chromosome 4, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_{π} (C, G), and S_{f} (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

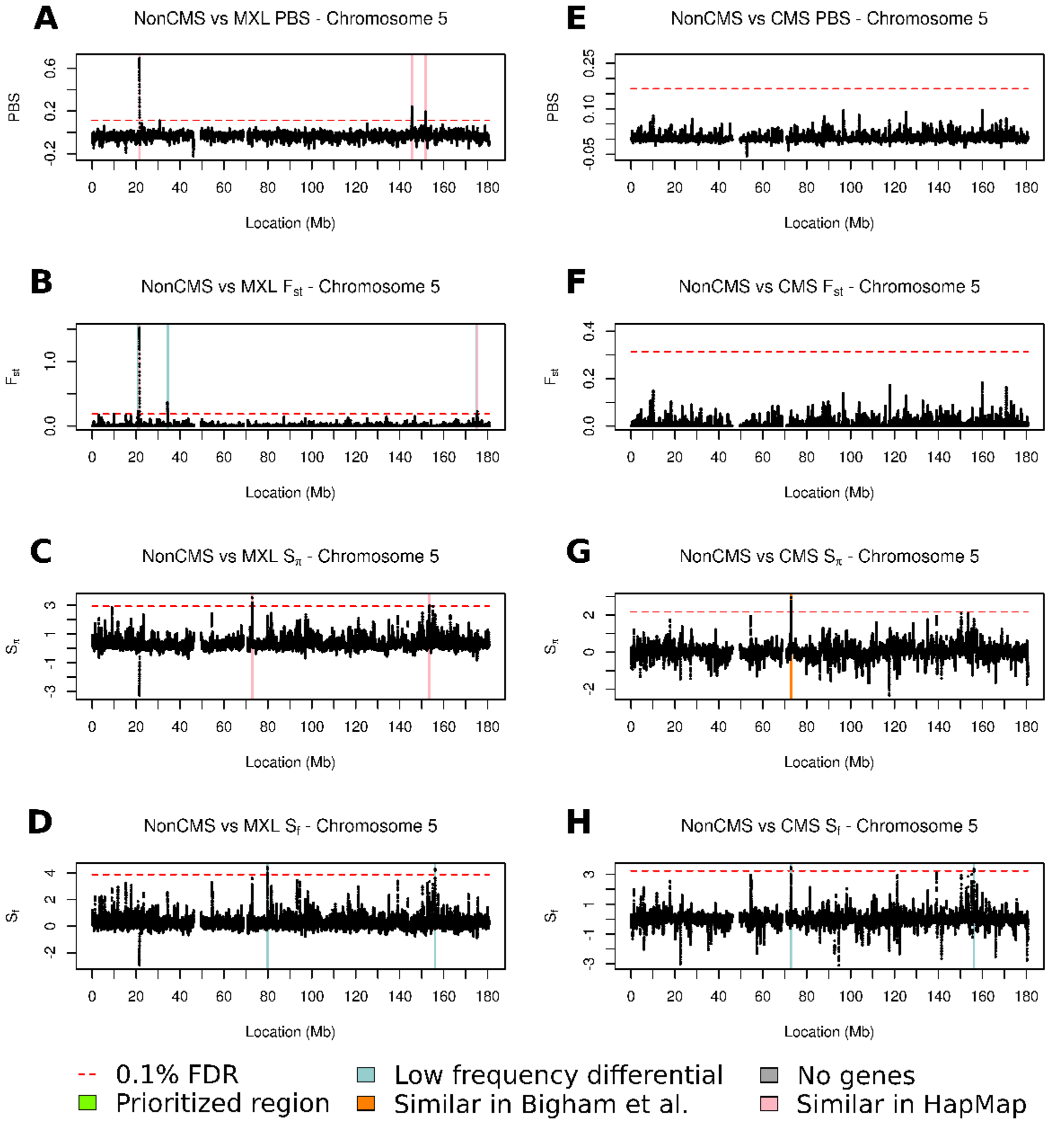


Figure S8. Test statistic values on chromosome 5, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_{π} (C, G), and S_{f} (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

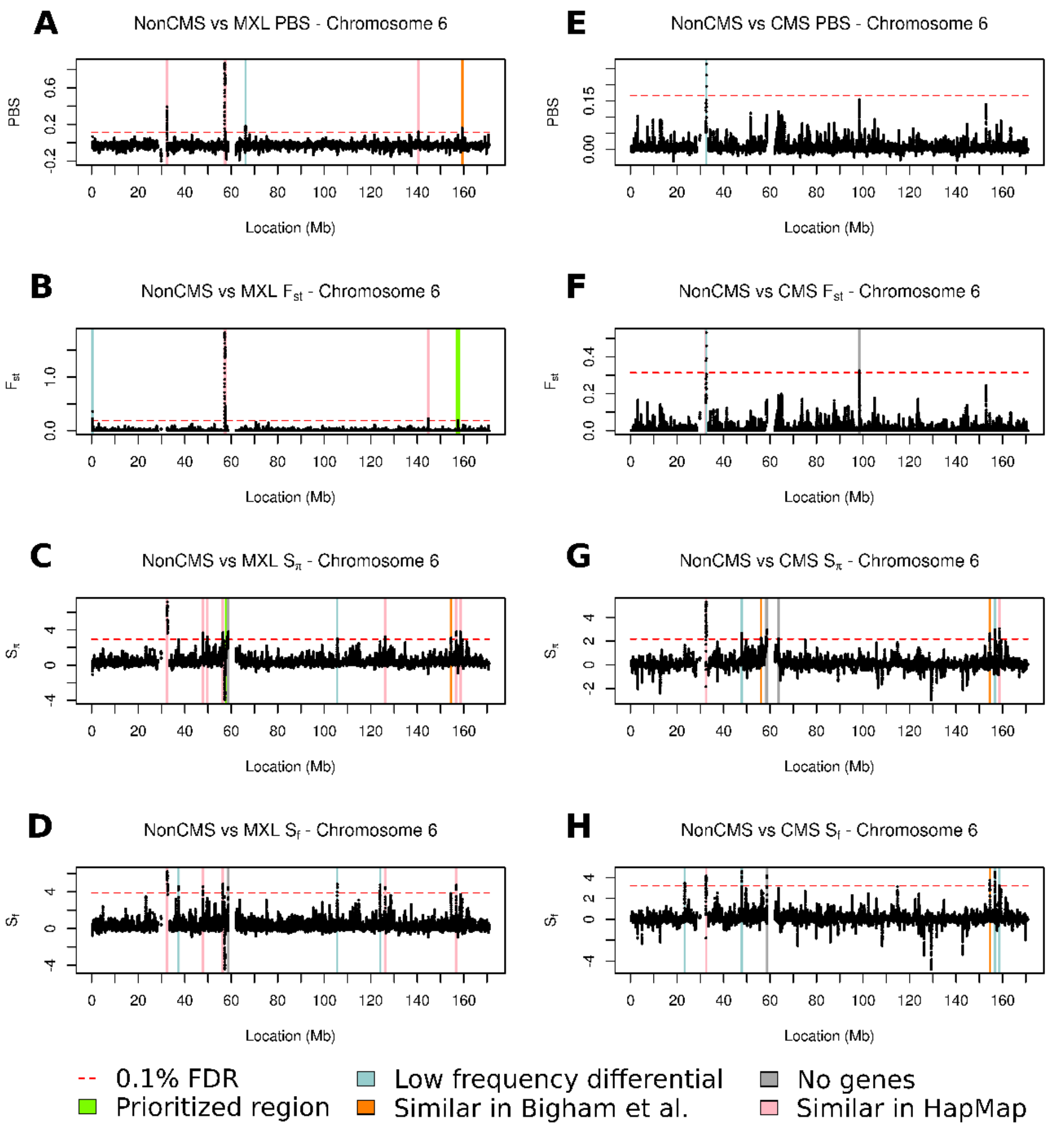


Figure S9. Test statistic values on chromosome 6, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in the non-CMS vs. MXL S_π statistic extends from positions 58,244,452-58,392,452, contains one gene (*GUSBP4*), and has maximum S_π statistic value of 3.238. The green region in the non-CMS vs. MXL F_{ST} statistic extends from positions 157,504,452-157,554,452, contains 1 gene (*ARID1B*) and has F_{ST} statistic value of 0.190.

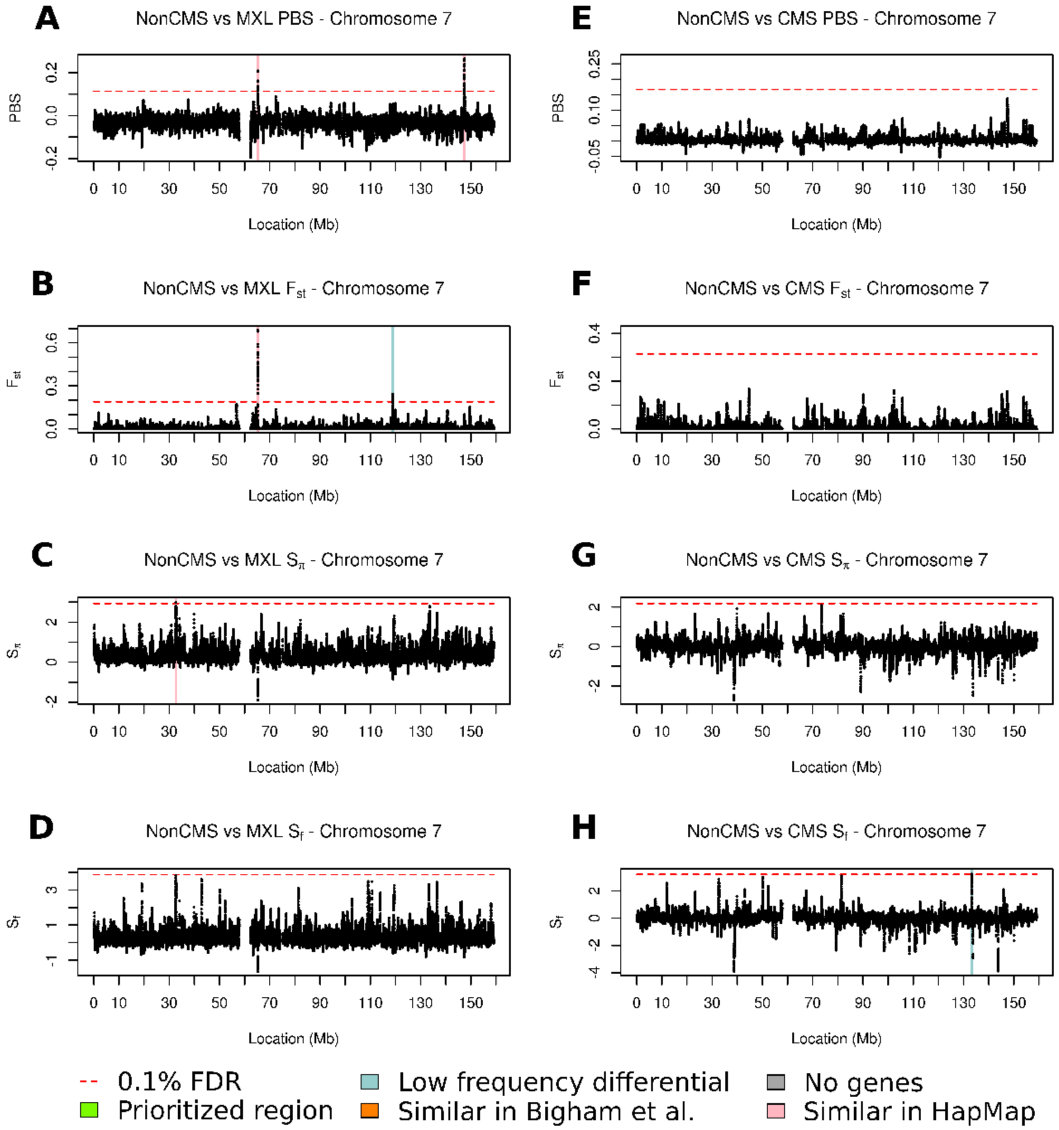


Figure S10. Test statistic values on chromosome 7, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

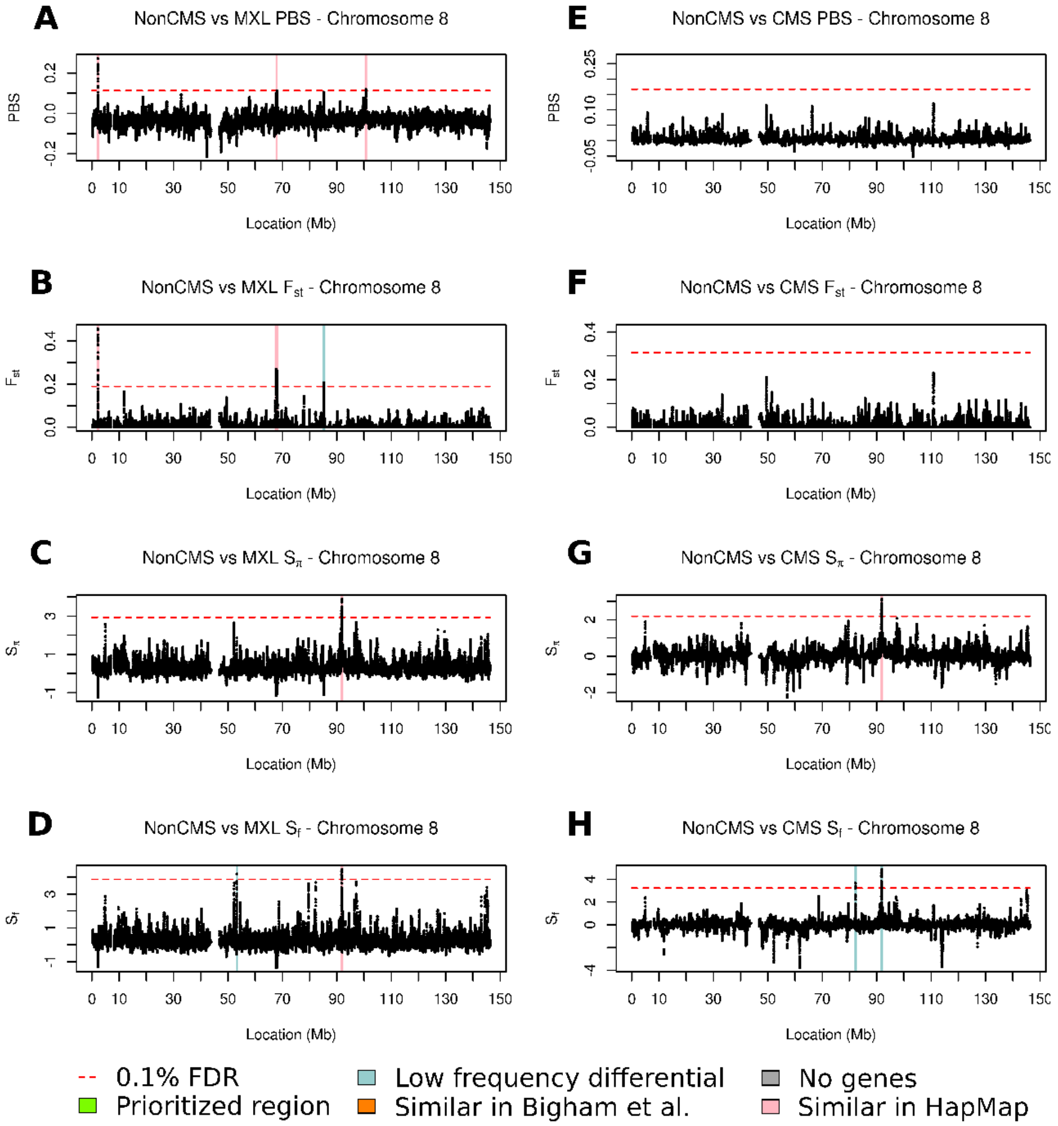


Figure S11. Test statistic values on chromosome 8, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

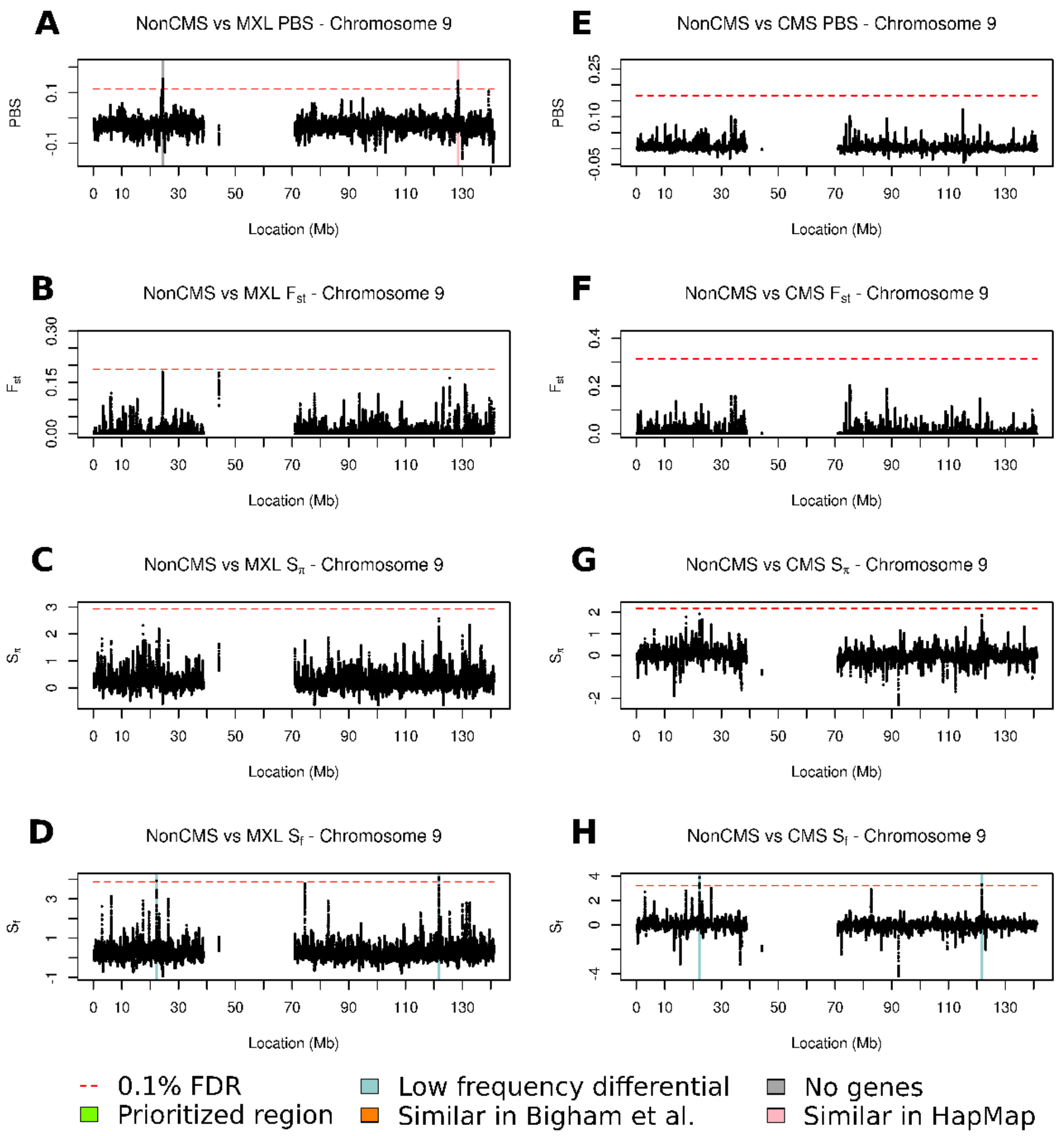


Figure S12. Test statistic values on chromosome 9, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

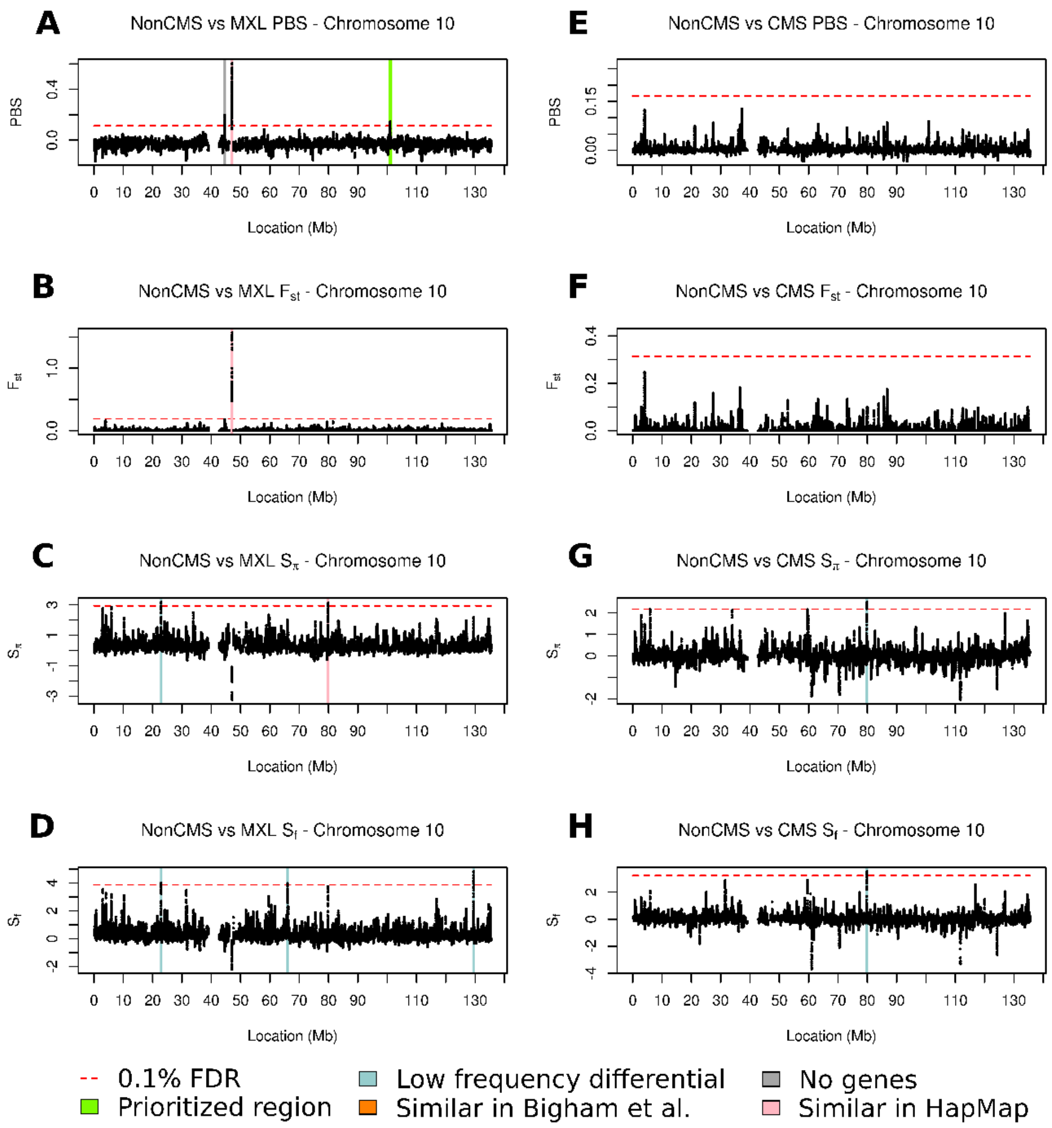


Figure S13. Test statistic values on chromosome 10, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. MXL *PBS* statistic extends from positions 101,014,523-101,092,523, contains 1 gene (*CNNM1*), and has maximum *PBS* statistic value of 0.146.

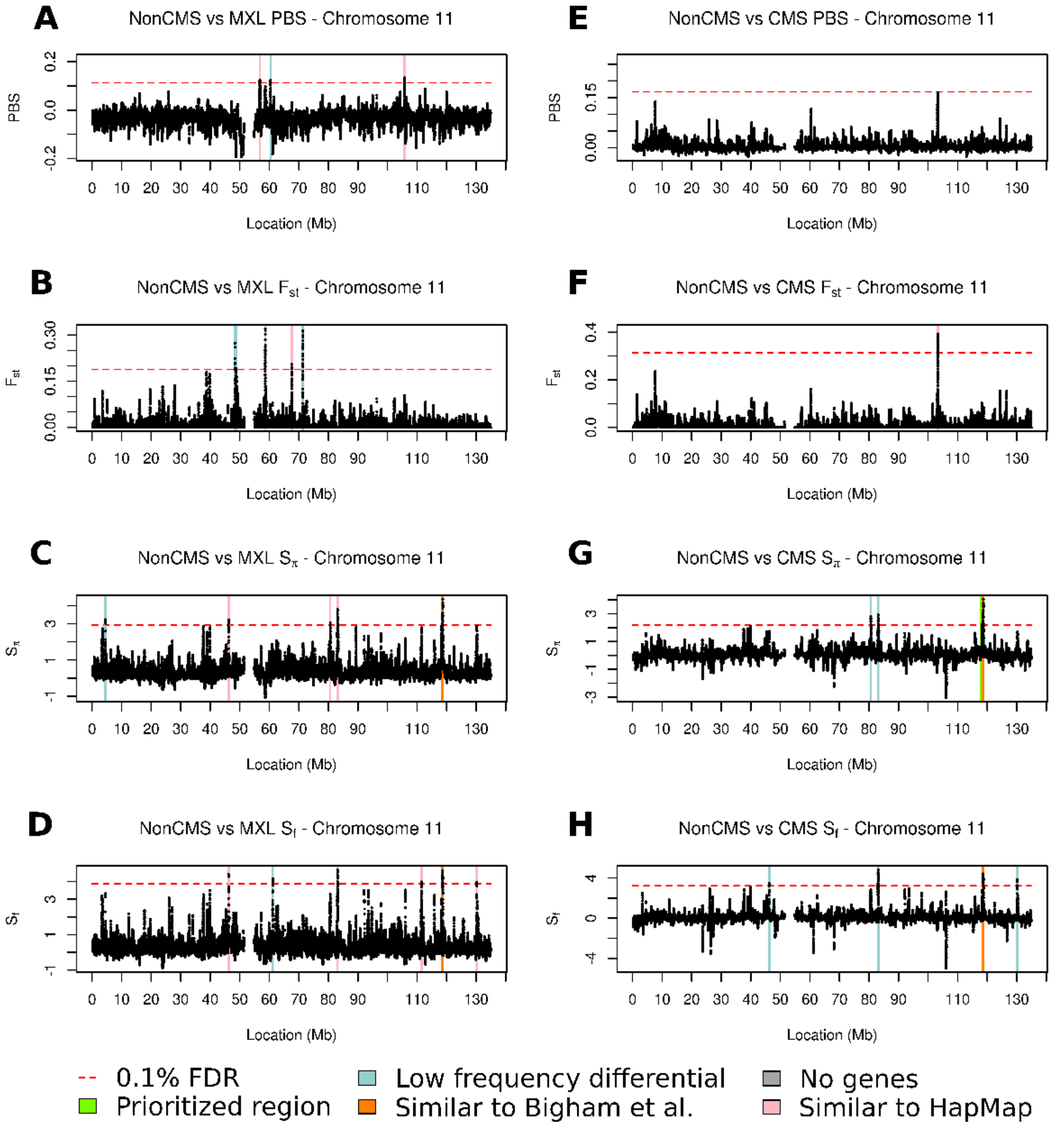


Figure S14. Test statistic values on chromosome 11, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. CMS S_π statistic extends from positions 118,147,948-118,199,948, contains 1 gene (*CD3E*), and has maximum S_π statistic value of 2.196.

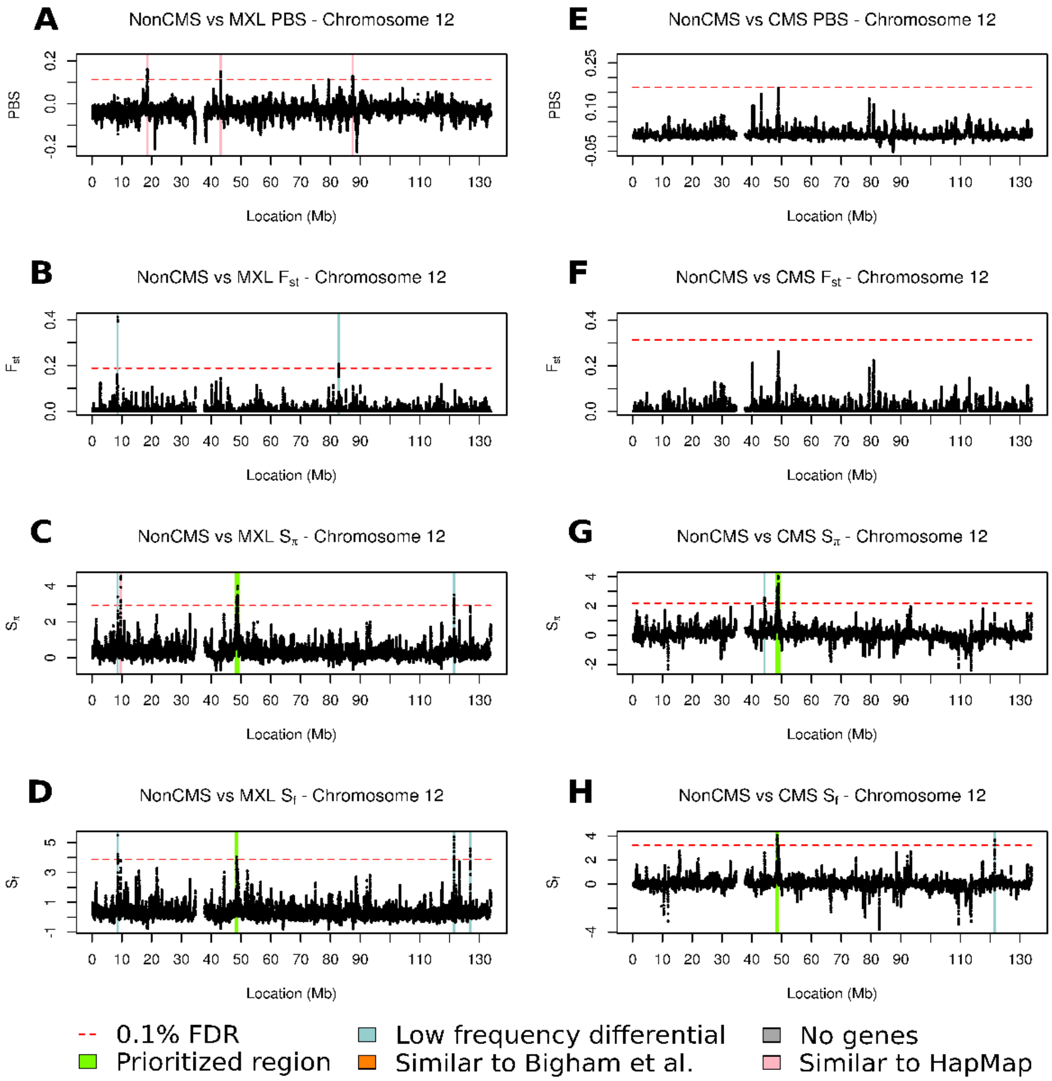


Figure S15. Test statistic values on chromosome 12, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green regions in non-CMS vs. CMS and non-CMS vs. MXL S_π and S_f statistics extend from positions 48,411,360-48,555,360, contain 3 genes (*SENP1*, *PFKM*, *ASB8*) and have maximum non-CMS vs. CMS S_π statistic value of 2.221, non-CMS vs. CMS S_f statistic value of 3.241, non-CMS vs. MXL S_π statistic value of 3.422, and non-CMS vs. MXL S_f statistic value of 4.051. The green regions in non-CMS vs. CMS and non-CMS vs. MXL S_π extend from 48,751,360 to 48,907,360, contain 2 genes (*ANP32D* and *C12orf54*), and have maximum non-CMS vs. CMS S_π statistic value of 2.212, and non-CMS vs. MXL statistic value of 4.007.

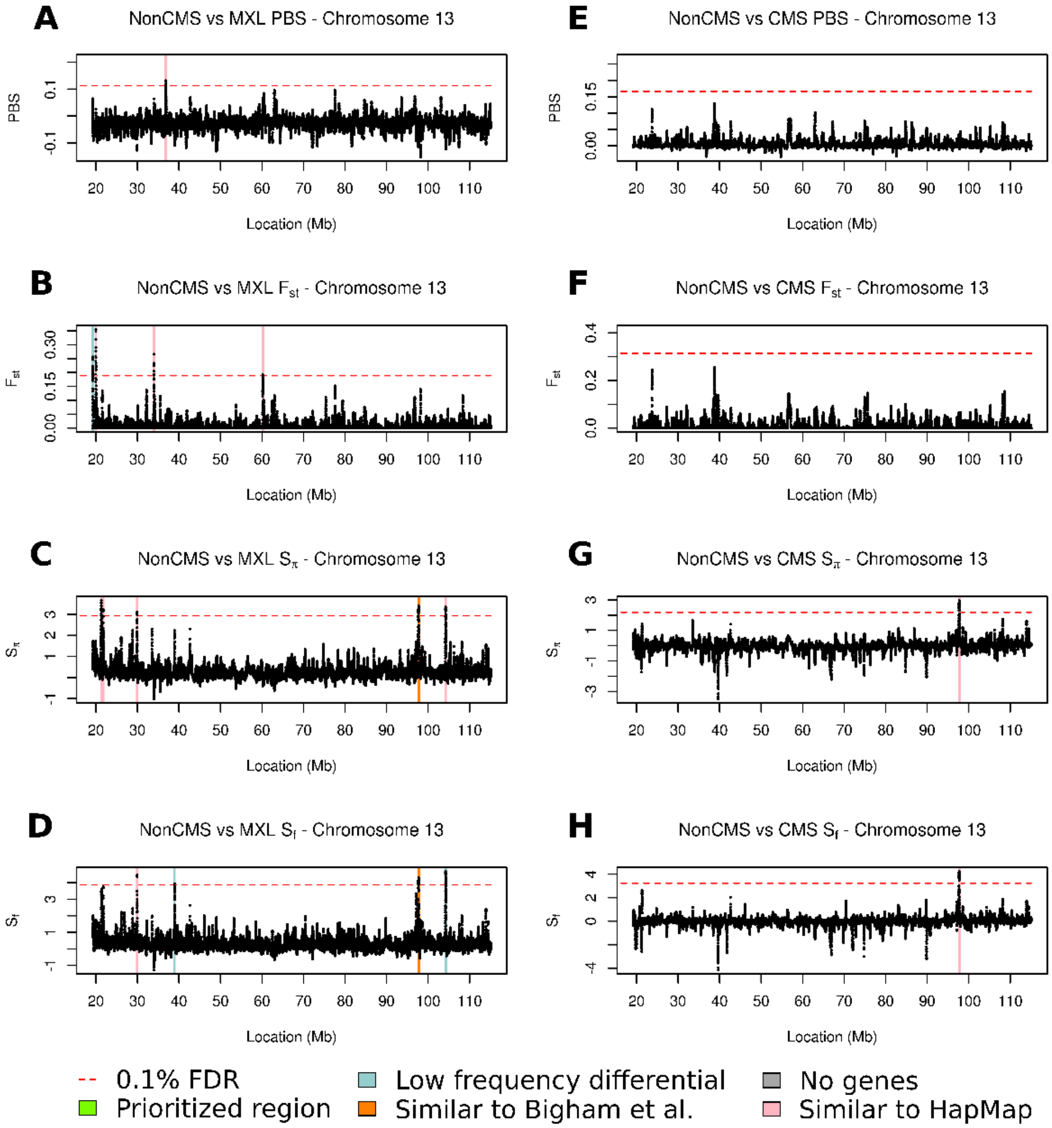


Figure S16. Test statistic values on chromosome 13, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_{f} (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

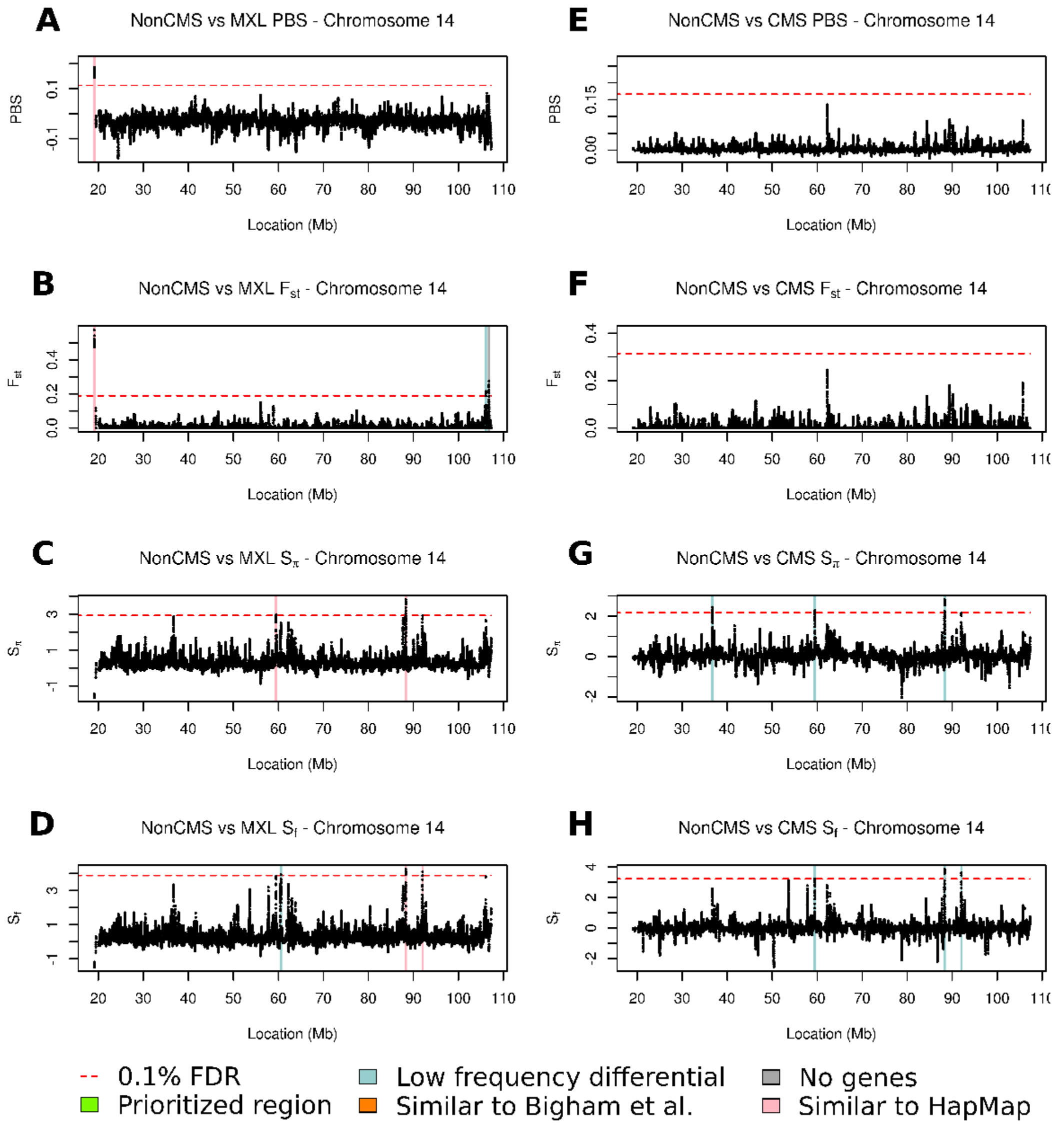


Figure S17. Test statistic values on chromosome 14, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_{π} (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

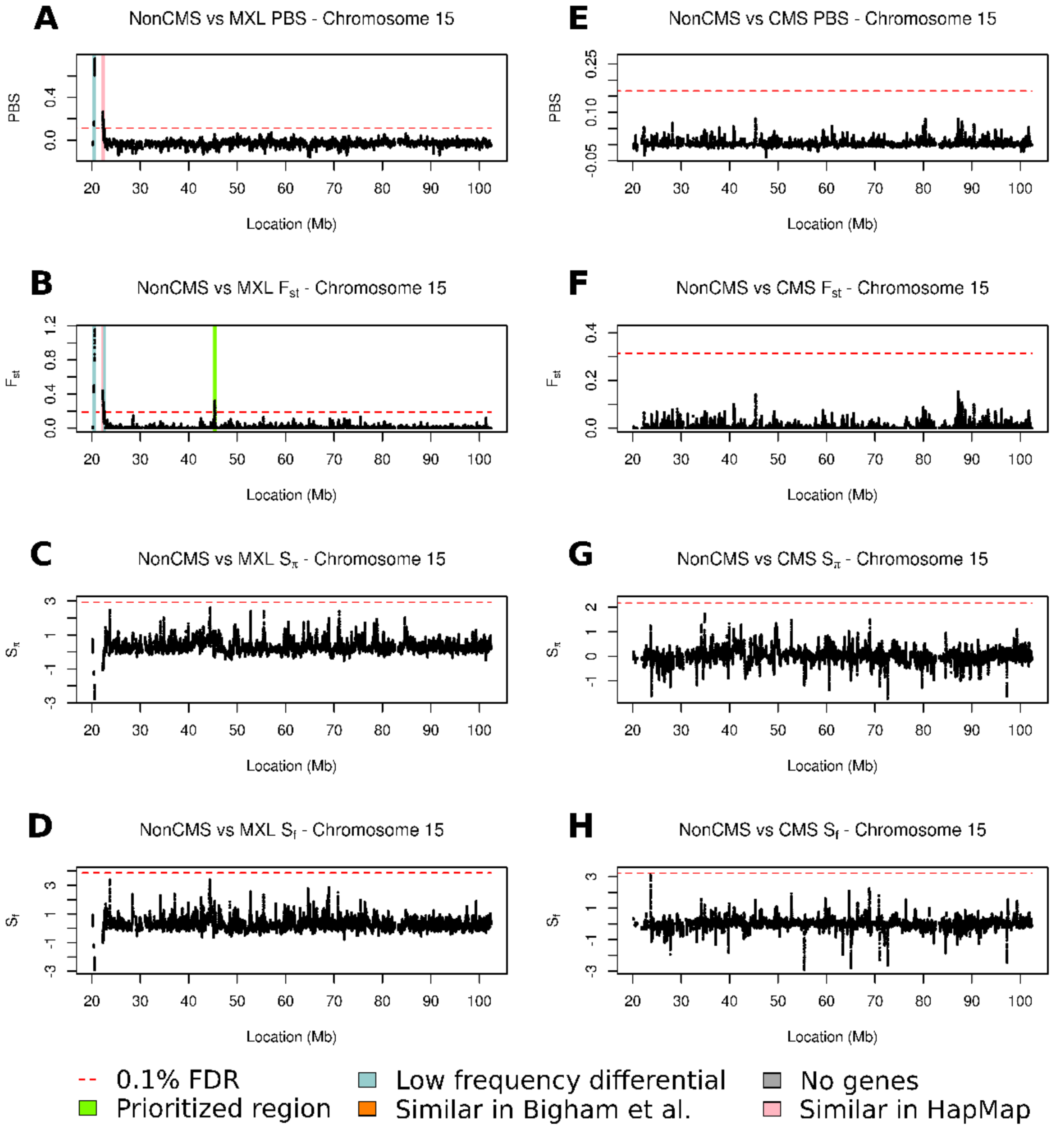


Figure S18. Test statistic values on chromosome 15, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. MXL F_{ST} statistic extends from positions 45,338,058-45,436,058, contains 5 genes (*SORD*, *DUOX2*, *DUOXA2*, *DUOXA1*, and *DUOX1*), and has maximum F_{ST} statistic value of 0.323.

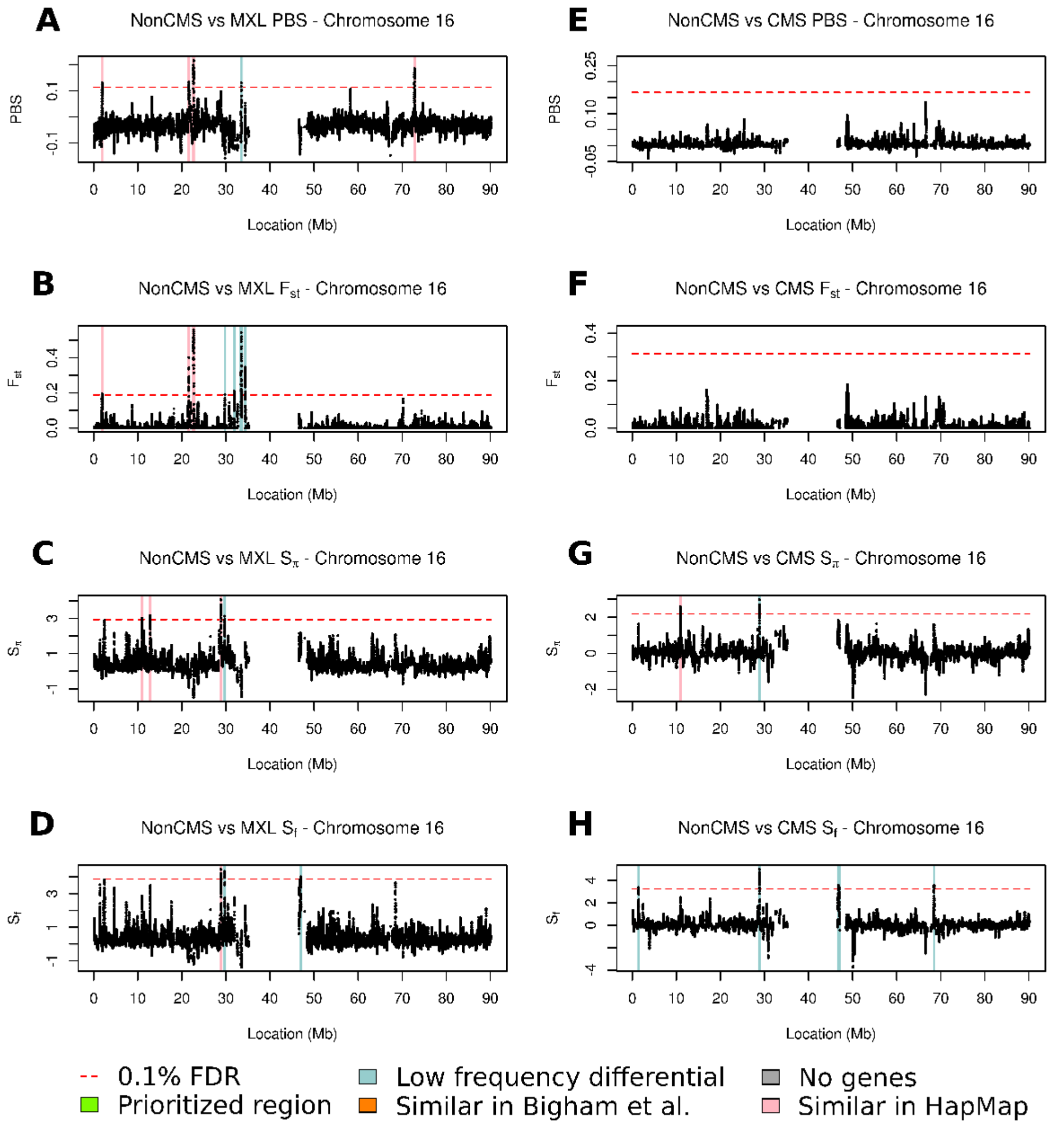


Figure S19. Test statistic values on chromosome 16, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_{π} (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

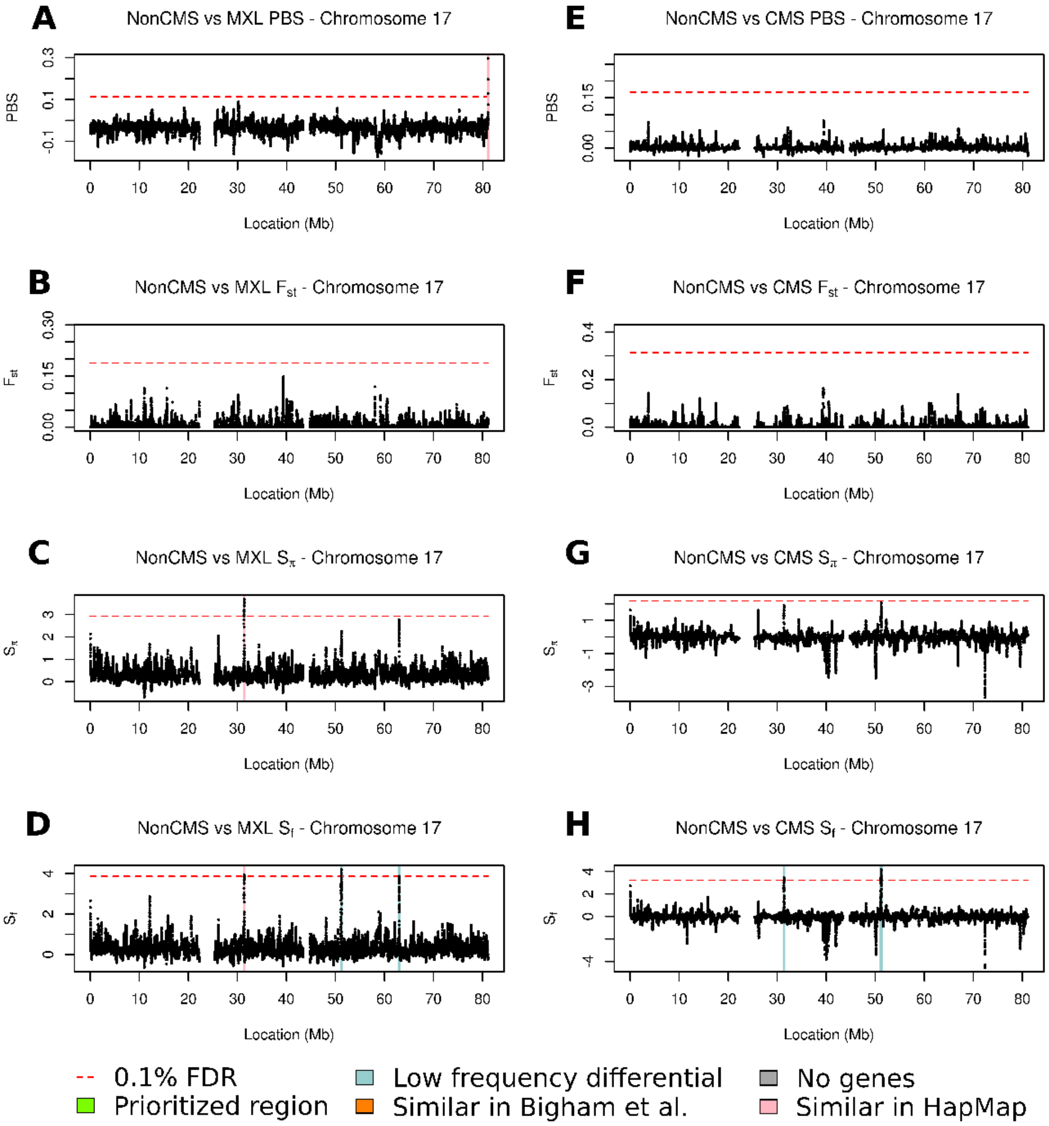


Figure S20. Test statistic values on chromosome 17, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

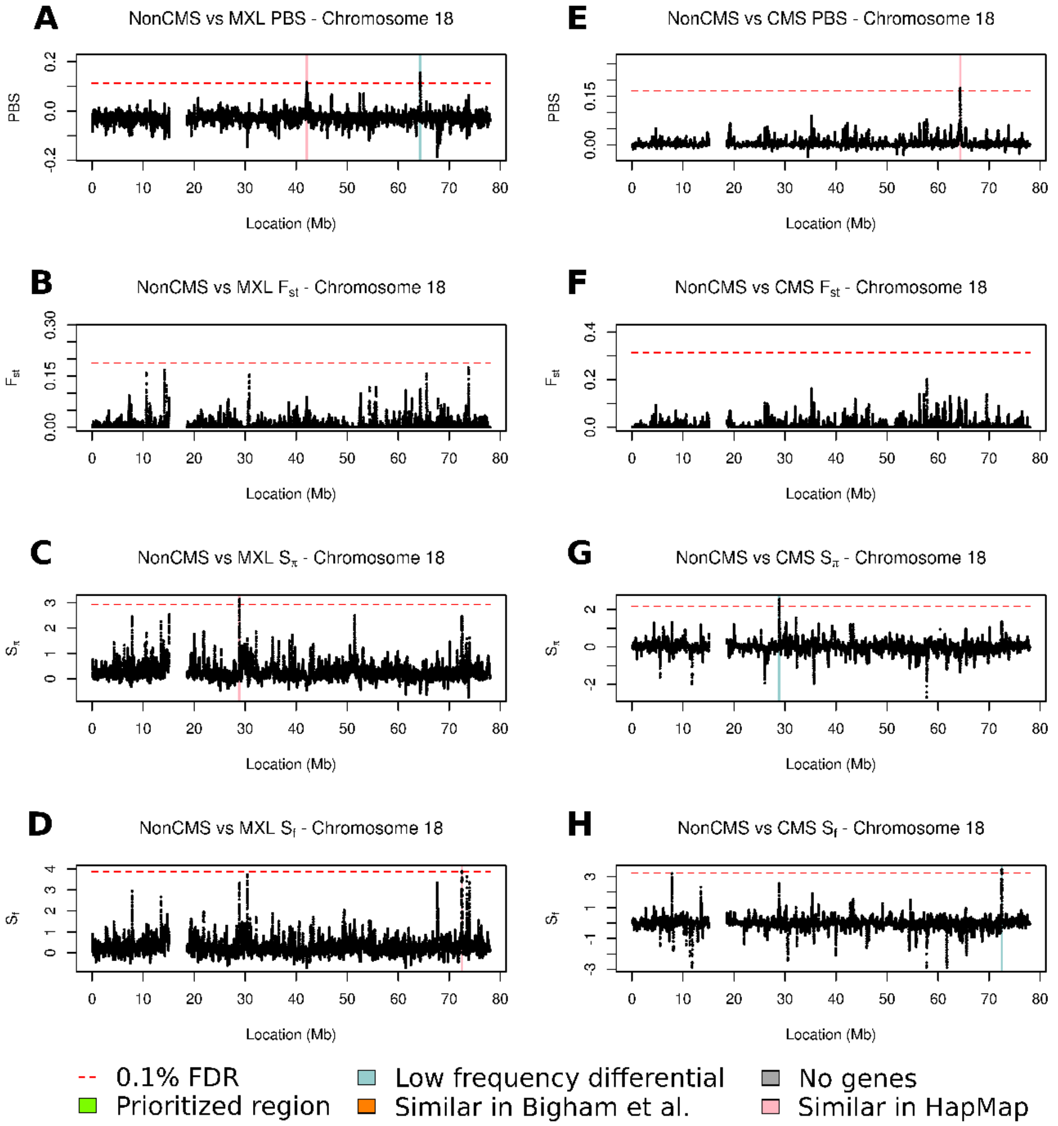


Figure S21. Test statistic values on chromosome 18, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

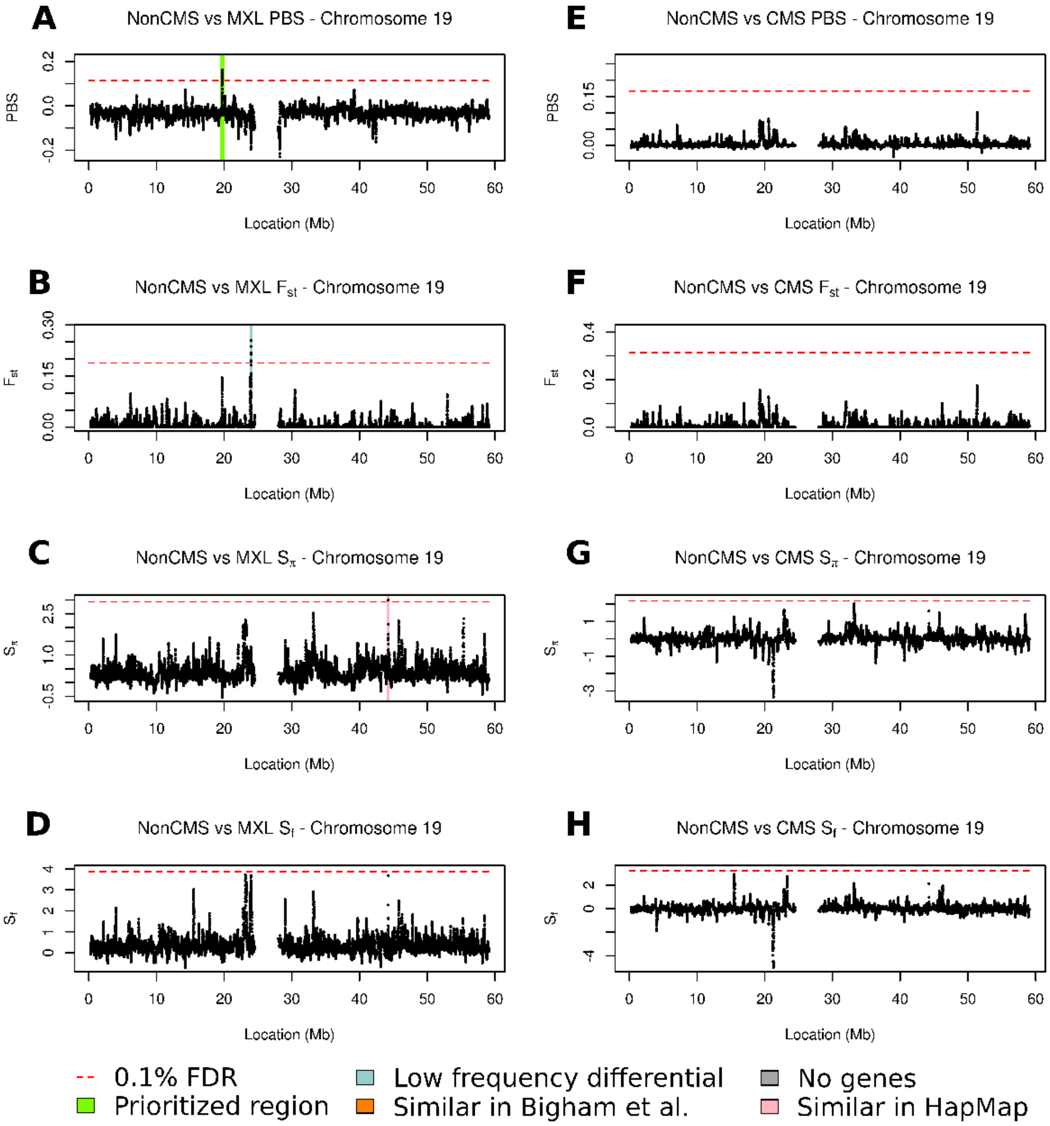


Figure S22. Test statistic values on chromosome 19, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. MXL *PBS* statistic extends from positions 19,665,844-19,747,844, contains 3 genes (*PBX4*, *LPAR2*, and *GMIP*), and has maximum *PBS* statistic value of 0.162.

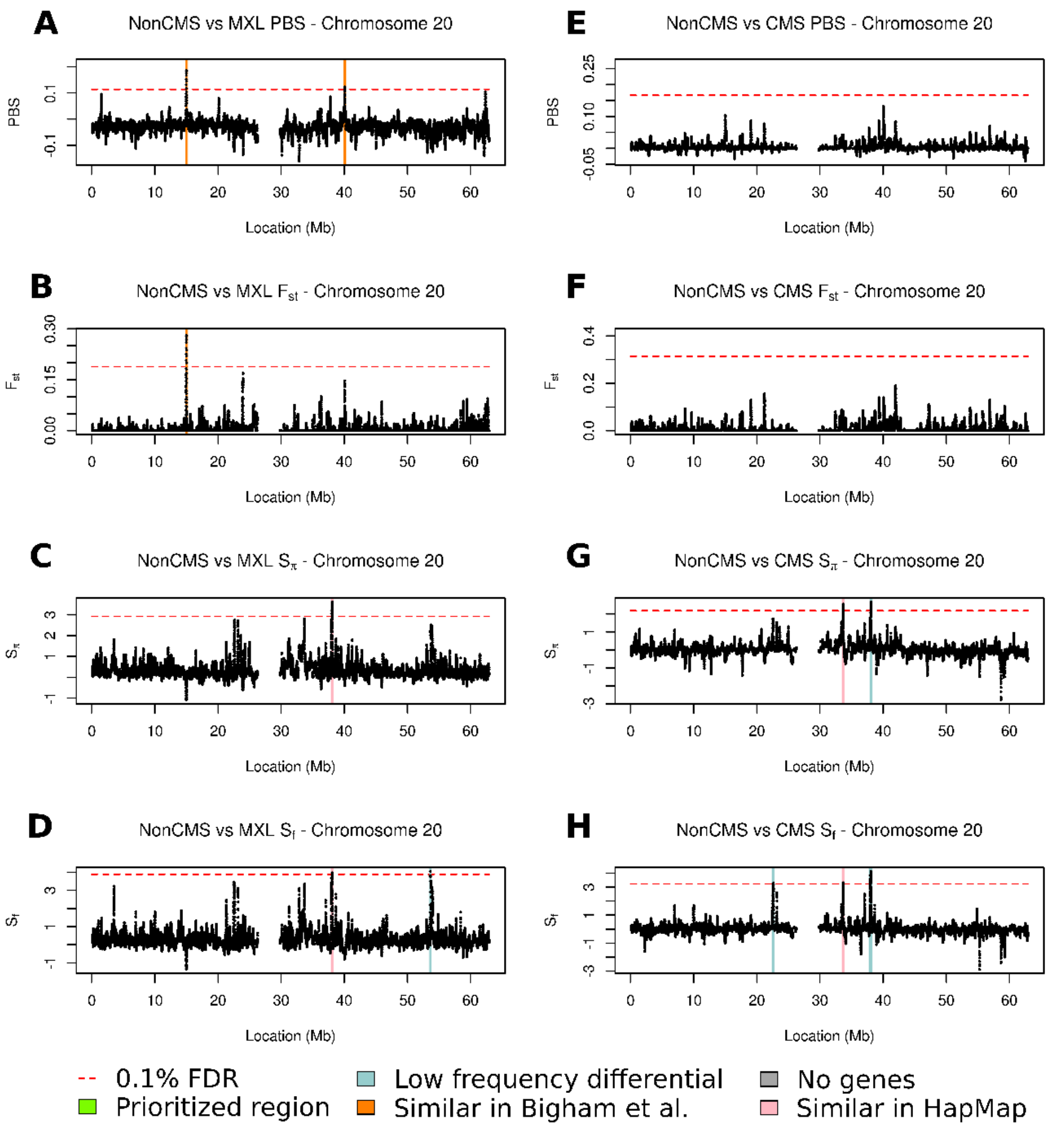


Figure S23. Test statistic values on chromosome 20, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are PBS (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

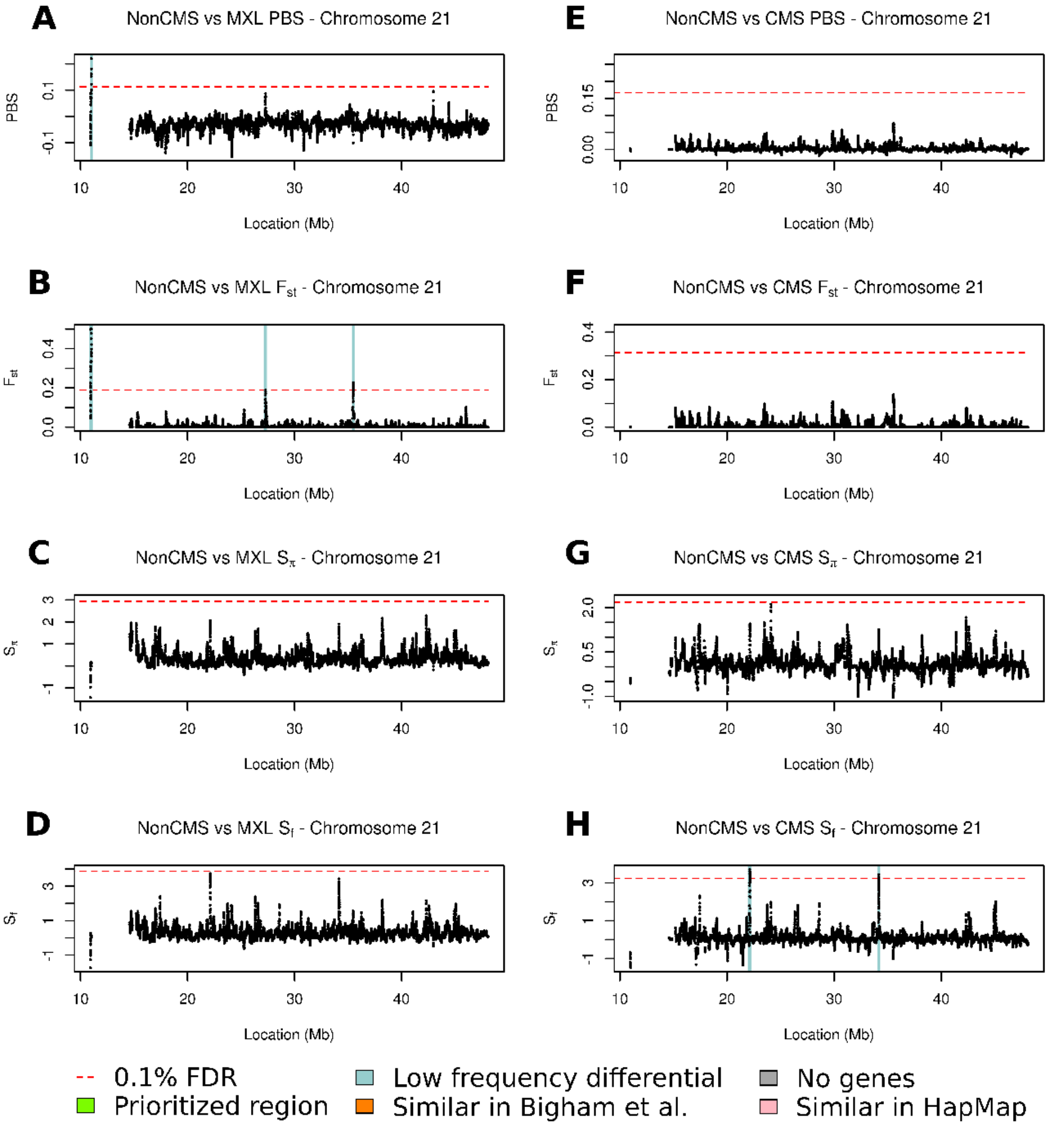


Figure S24. Test statistic values on chromosome 21, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

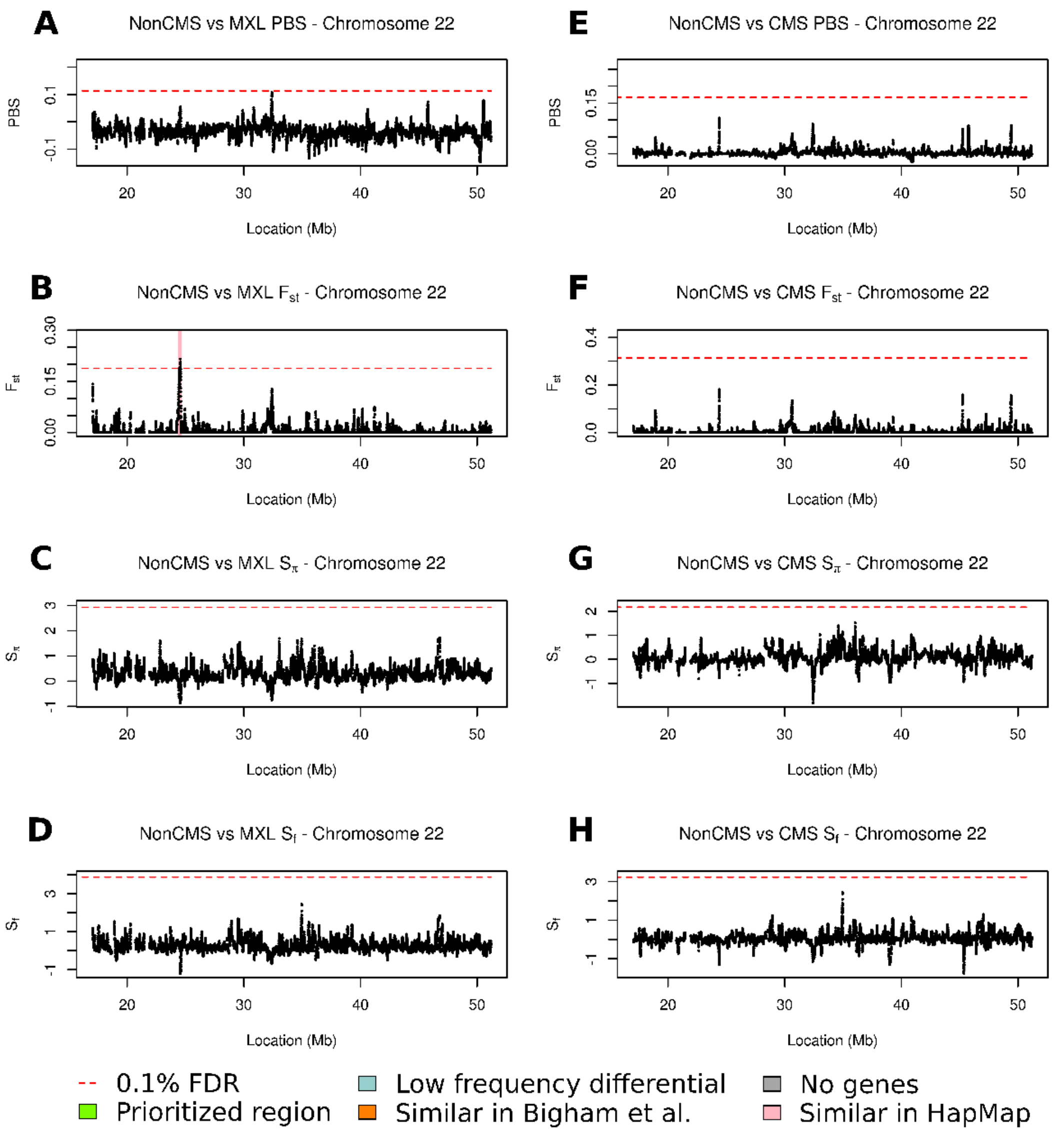


Figure S25. Test statistic values on chromosome 22, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green.

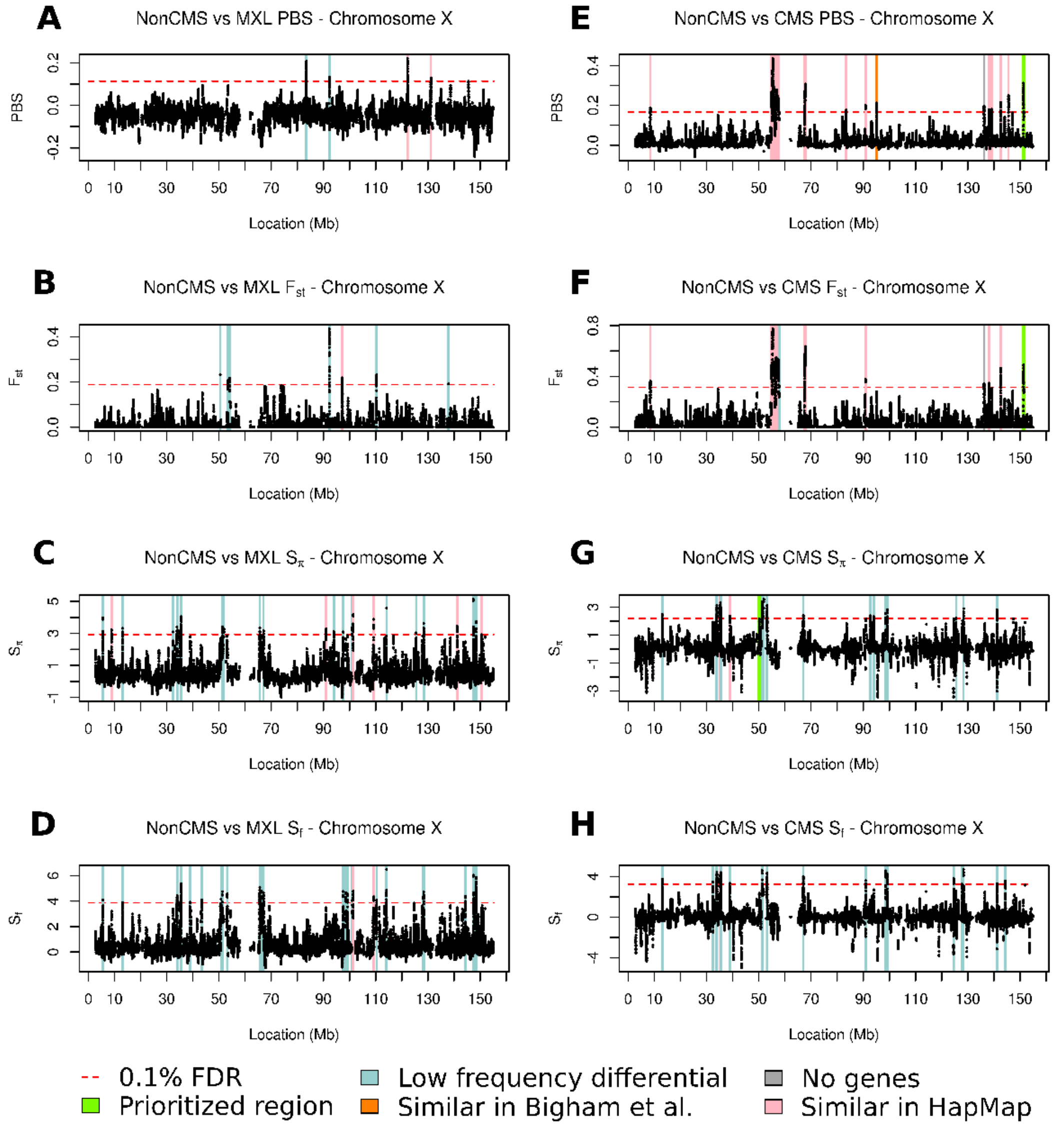
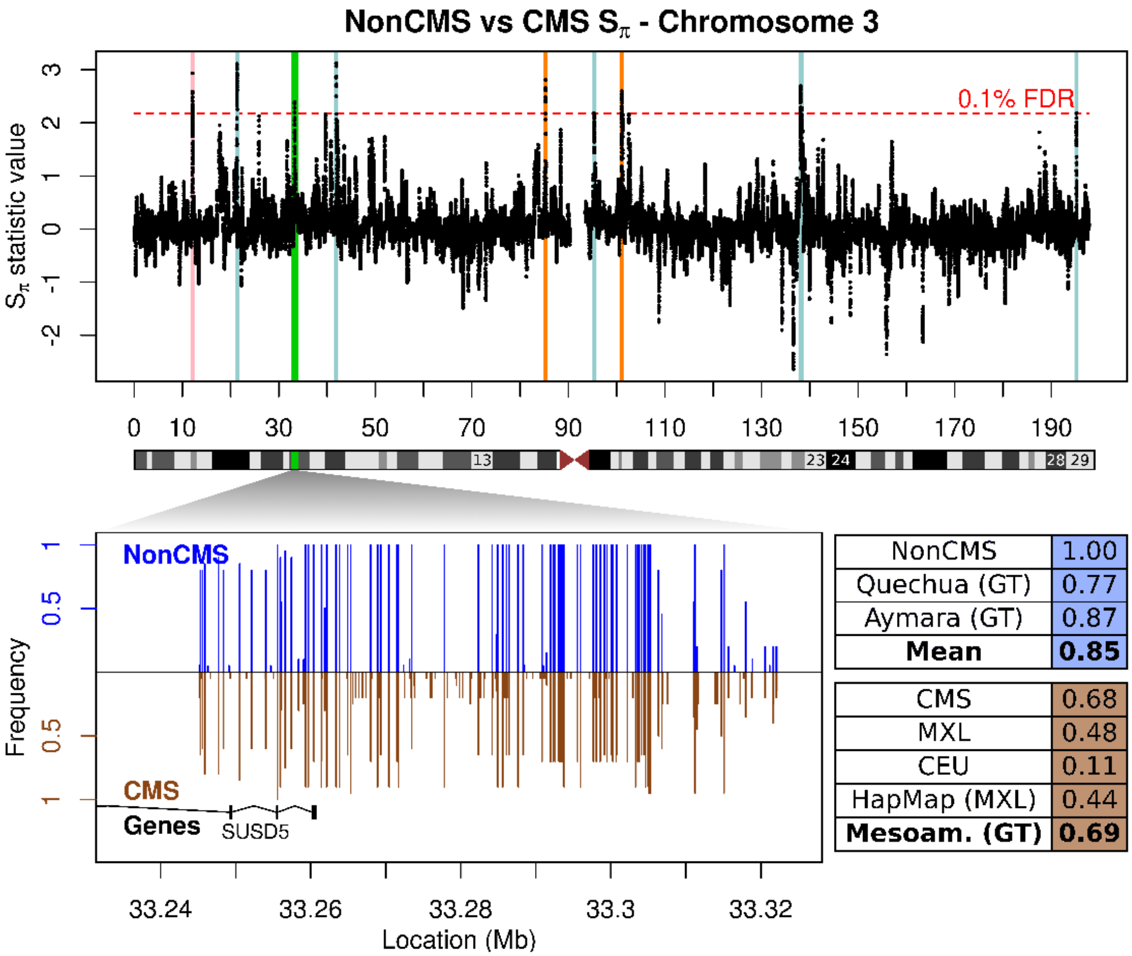


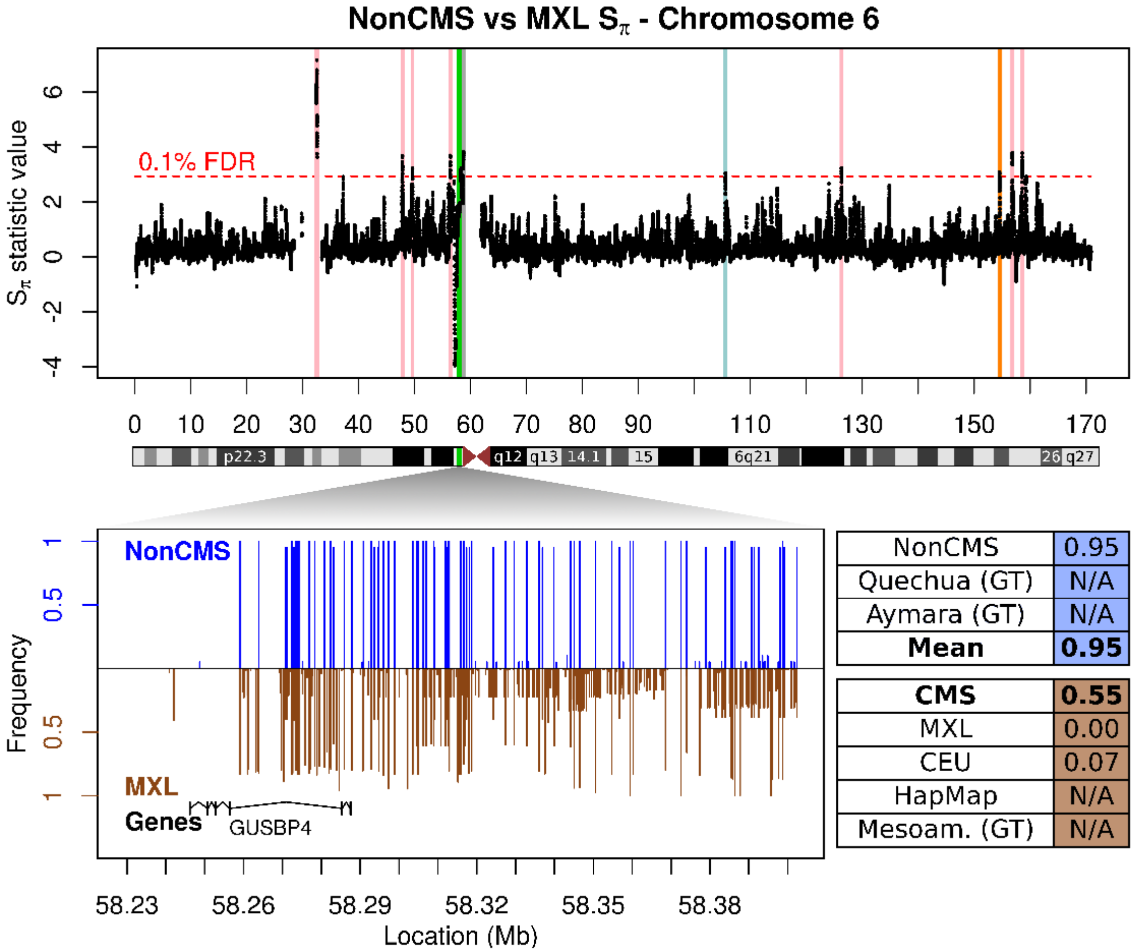
Figure S26. Test statistic values on chromosome X, when applied to non-CMS vs. MXL (A-D) and to non-CMS vs. CMS (E-H). Tests shown are *PBS* (A, E), F_{ST} (B, F), S_π (C, G), and S_f (D, H). The prioritization criteria used to shortlist the regions are color-coded. Regions above the 0.1% FDR that made it to the final list of prioritized regions are green. Specifically, the green region in non-CMS vs. CMS S_π statistic extends from positions 50,147,676-50,197,676, contains 1 gene (*DGKK*), and has maximum S_π statistic value of 2.179. The green region in non-CMS vs. CMS *PBS* and F_{ST} statistics extends from position 151,275,676-151,421,676, contains 3 genes (*MAGEA5*, *MAGEA10*, and *GABRA3*), and has maximum *PBS* statistic value of 0.185, and maximum F_{ST} statistic value of 0.318.



NonCMS	1.00
Quechua (GT)	0.77
Aymara (GT)	0.87
Mean	**0.85**

CMS	0.68
MXL	0.48
CEU	0.11
HapMap (MXL)	0.44
Mesoam. (GT)	**0.69**

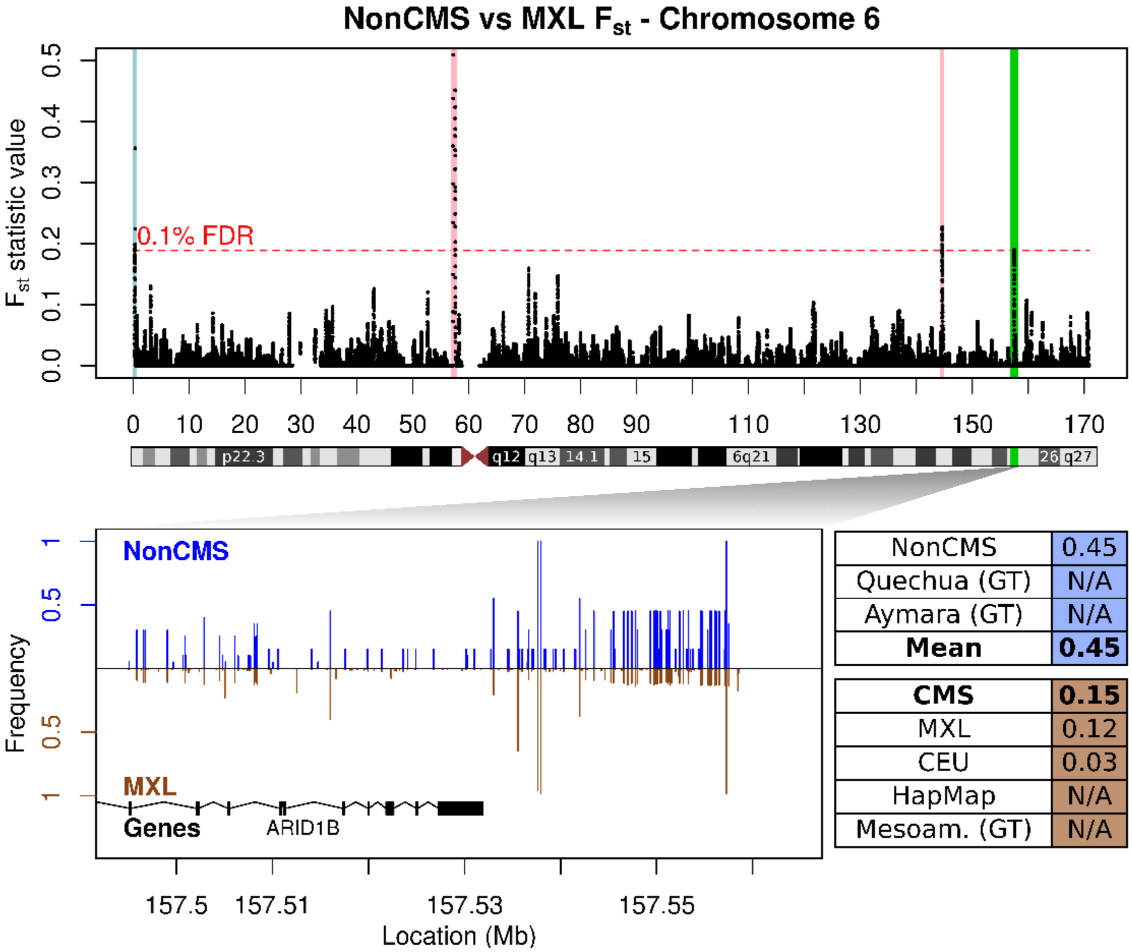
Figure S27. Profile of the chromosome 3 candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\mathrm{CMS}}$) plotted across chromosome 3. Ten distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 3p22.3). As can be seen, there is an almost complete fixation of a haplotype in the non-CMS population that is at a much lower frequency in all lowlander and mal-adapted controls.



NonCMS	0.95
Quechua (GT)	N/A
Aymara (GT)	N/A
Mean	**0.95**

CMS	**0.55**
MXL	0.00
CEU	0.07
HapMap	N/A
Mesoam. (GT)	N/A

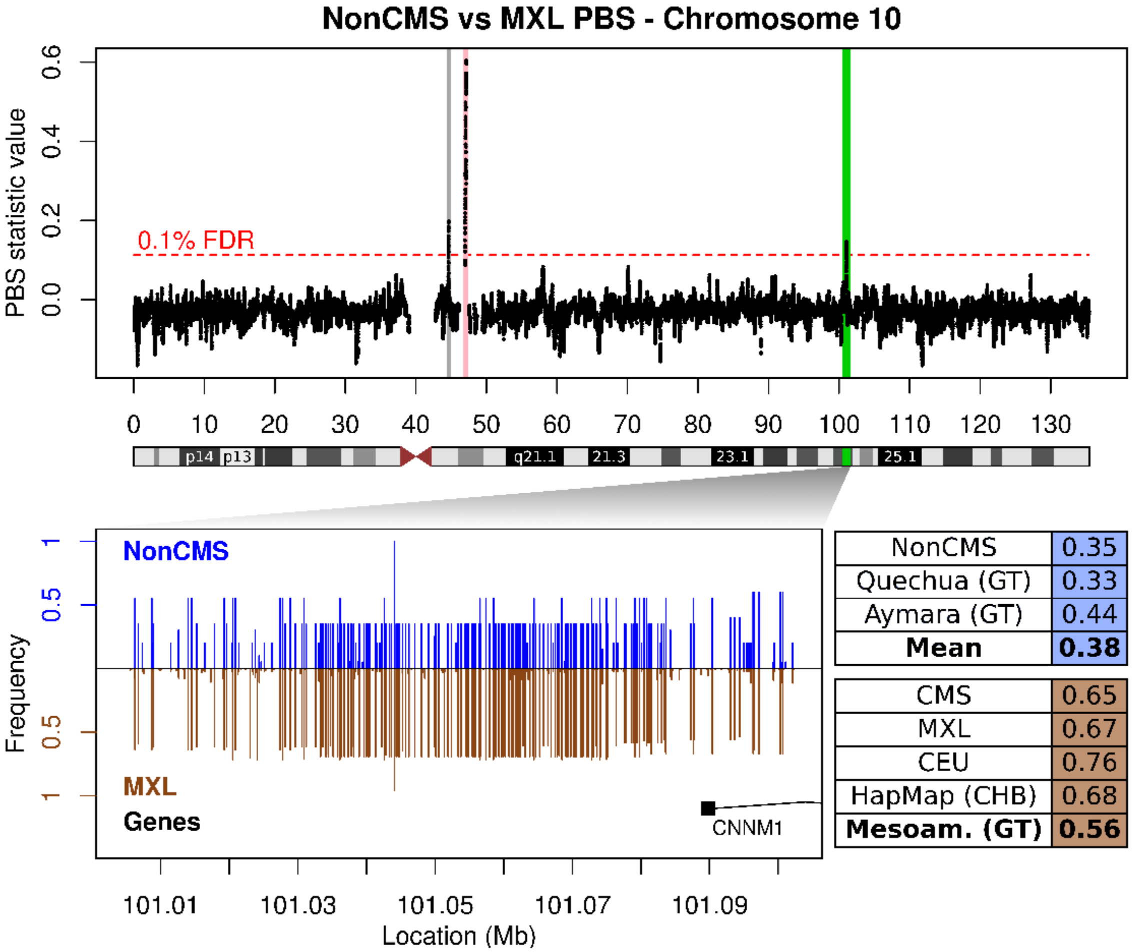
Figure S28. Profile of the chromosome 6 (S_π) candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\mathrm{MXL}}$) plotted across chromosome 6. Eleven distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 6p11.2). As can be seen, there is an almost complete fixation of a haplotype in the non-CMS population that is at a much lower frequency in all sampled lowlander and mal-adapted controls.



NonCMS	0.45
Quechua (GT)	N/A
Aymara (GT)	N/A
Mean	**0.45**

CMS	**0.15**
MXL	0.12
CEU	0.03
HapMap	N/A
Mesoam. (GT)	N/A

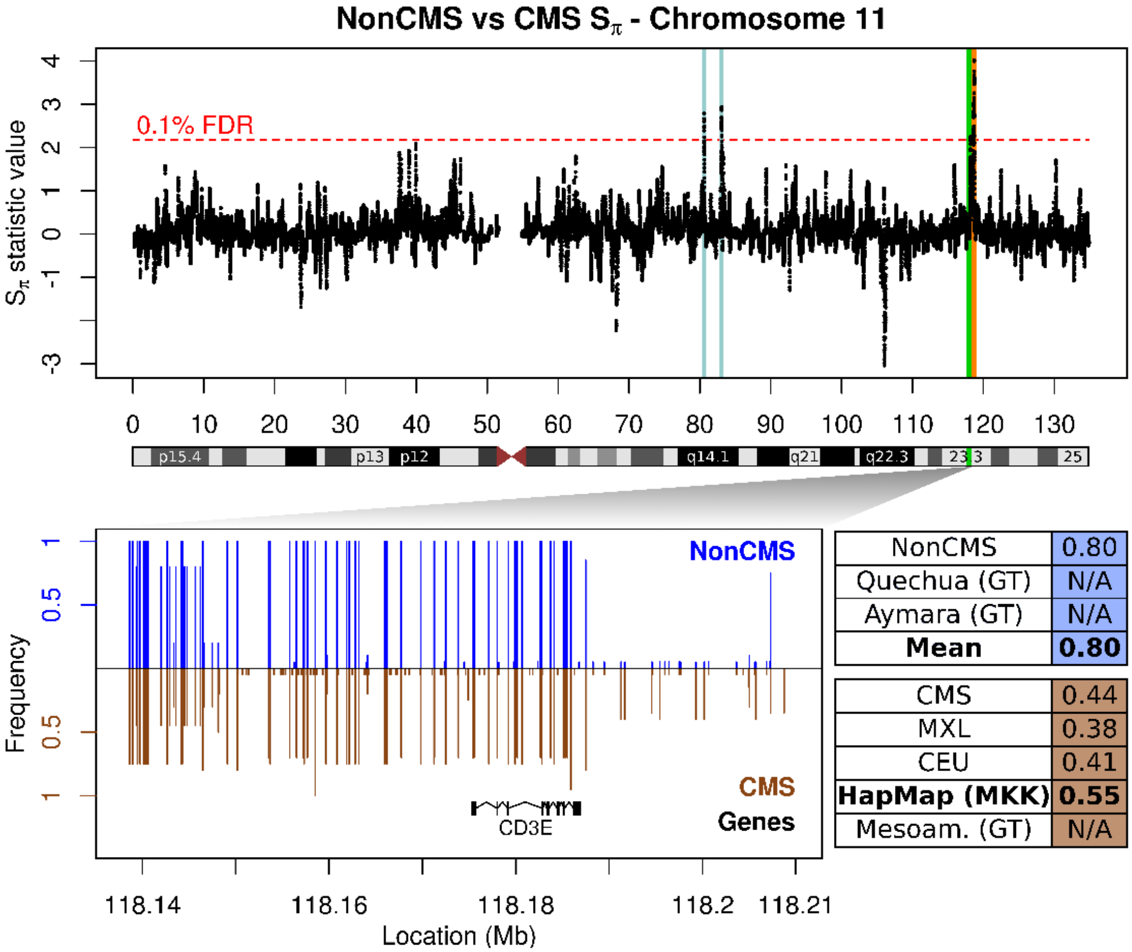
Figure S29. Profile of the chromosome 6 (F_{ST}) candidate region. (Top) The statistic under which this region is significant ($F_{\mathrm{ST,MXL}}$) plotted across chromosome 6. Six distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 6q25.3). As can be seen, the haplotype in this region is at a much higher frequency in the non-CMS population relative to all sampled lowlander and mal-adapted controls.



NonCMS	0.35
Quechua (GT)	0.33
Aymara (GT)	0.44
Mean	**0.38**

CMS	0.65
MXL	0.67
CEU	0.76
HapMap (CHB)	0.68
Mesoam. (GT)	**0.56**

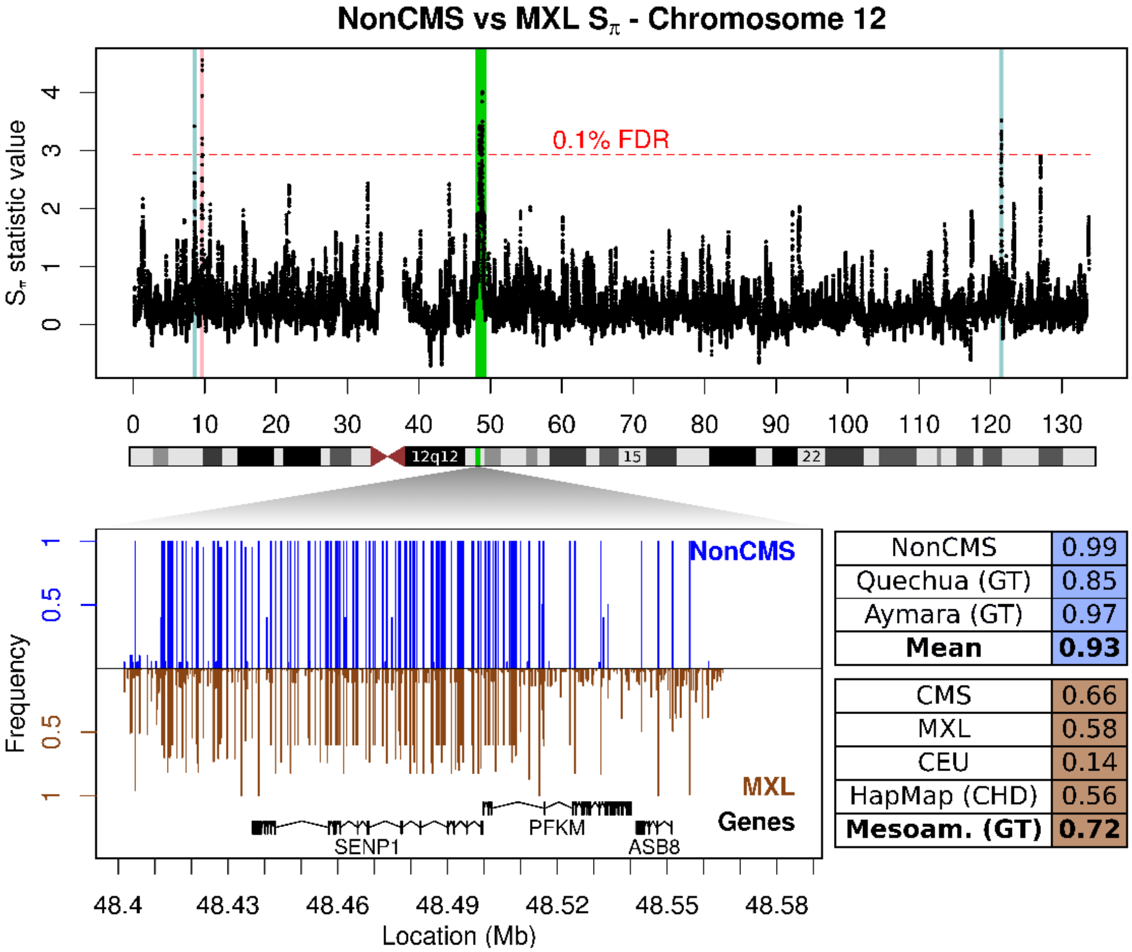
Figure S30. Profile of the chromosome 10 (*PBS*) candidate region. (Top) The statistic under which this region is significant (PBS_{MXL}) plotted across chromosome 10. Three distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 10q24.2). As can be seen, the reference haplotype in this region is at a much lower frequency in the non-CMS population relative to all sampled lowlander and mal-adapted controls.



NonCMS	0.80
Quechua (GT)	N/A
Aymara (GT)	N/A
Mean	**0.80**

CMS	0.44
MXL	0.38
CEU	0.41
HapMap (MKK)	**0.55**
Mesoam. (GT)	N/A

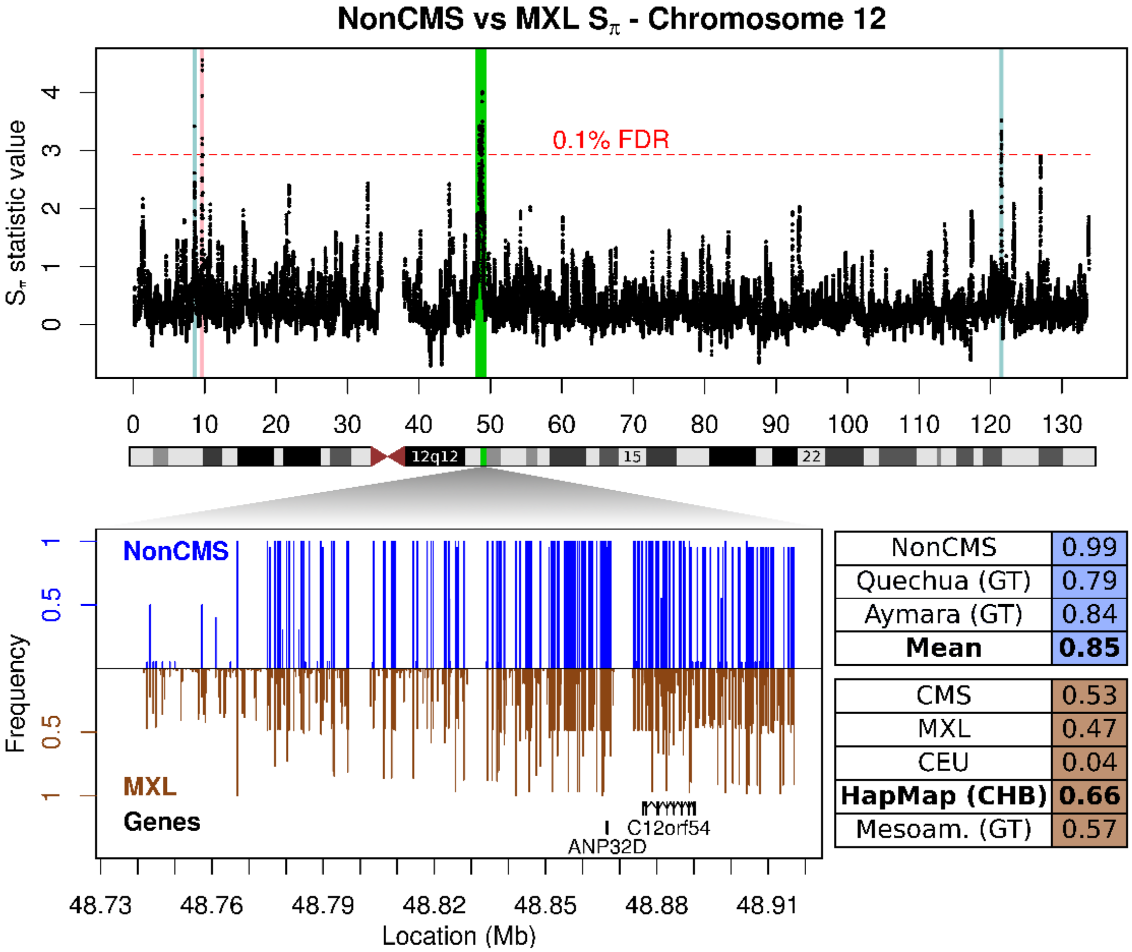
Figure S31. Profile of the chromosome 11 (S_π) candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\mathrm{CMS}}$) plotted across chromosome 11. Four distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 11q23.3). As can be seen, the haplotype in this region is at a much higher frequency in the non-CMS population relative to all sampled lowlander and mal-adapted controls.



NonCMS	0.99
Quechua (GT)	0.85
Aymara (GT)	0.97
Mean	**0.93**

CMS	0.66
MXL	0.58
CEU	0.14
HapMap (CHD)	0.56
Mesoam. (GT)	**0.72**

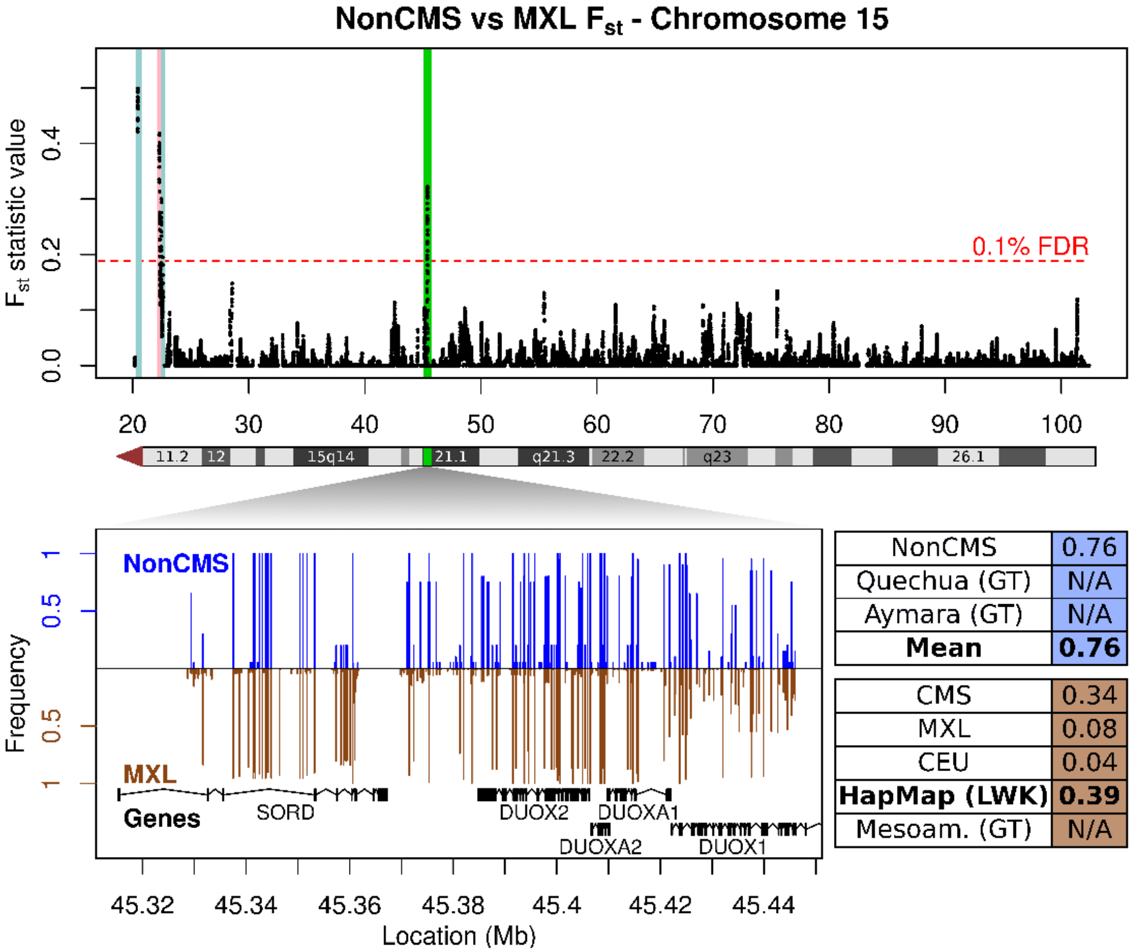
Figure S32. Profile of the first chromosome 12 (S_π) candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\mathrm{MXL}}$) plotted across chromosome 12. Five distinct regions exceeded the 0.1% FDR threshold, out of which two were considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 12q13.11). As can be seen, there is an almost complete fixation of a haplotype in the non-CMS population that is at a much lower frequency in all sampled lowlander and mal-adapted controls.



NonCMS	0.99
Quechua (GT)	0.79
Aymara (GT)	0.84
Mean	**0.85**

CMS	0.53
MXL	0.47
CEU	0.04
HapMap (CHB)	**0.66**
Mesoam. (GT)	0.57

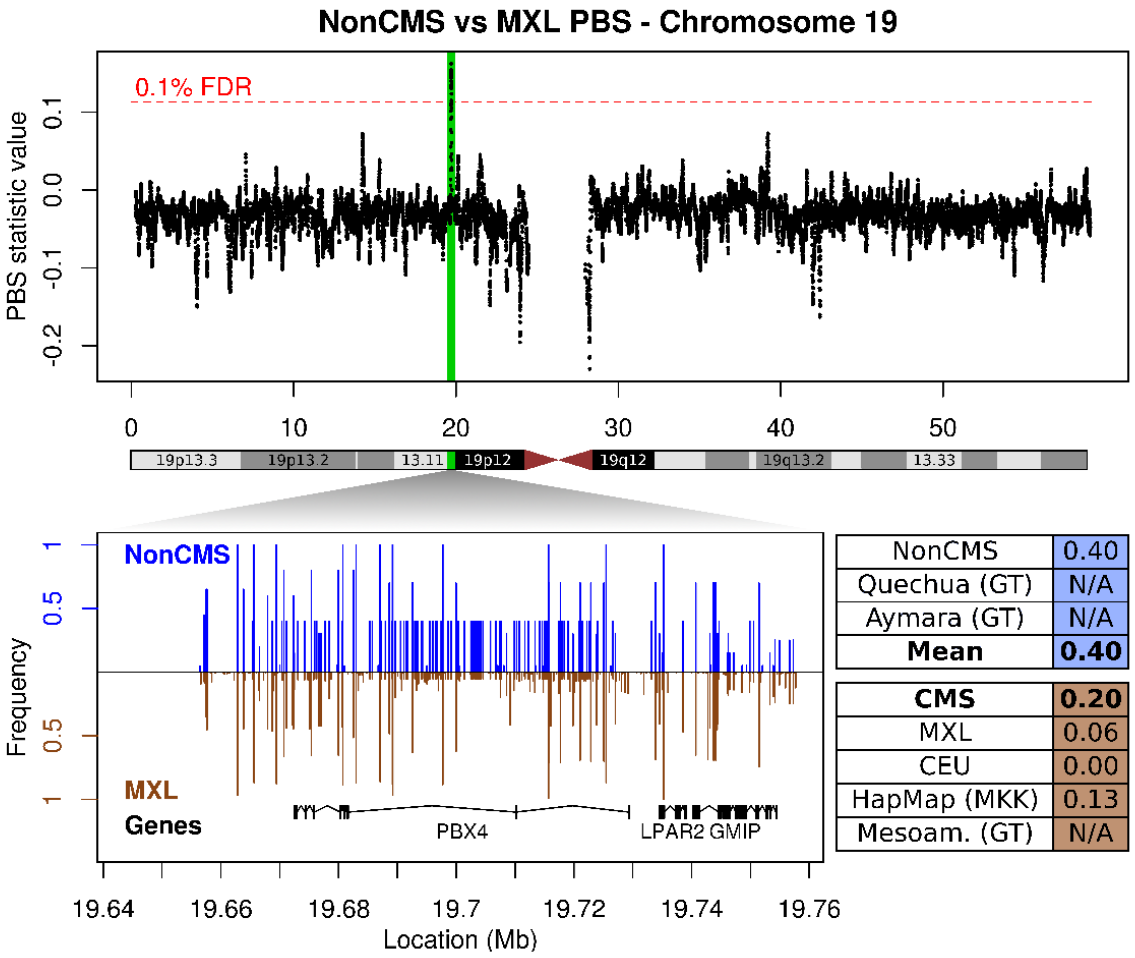
Figure S33. Profile of the second chromosome 12 (S_π) candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\mathrm{MXL}}$) plotted across chromosome 12. Five distinct regions exceeded the 0.1% FDR threshold, out of which two were considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 12q13.11). As can be seen, there is an almost complete fixation of a haplotype in the non-CMS population that is at a much lower frequency in all sampled lowlander and mal-adapted controls.



NonCMS	0.76
Quechua (GT)	N/A
Aymara (GT)	N/A
Mean	**0.76**

CMS	0.34
MXL	0.08
CEU	0.04
HapMap (LWK)	**0.39**
Mesoam. (GT)	N/A

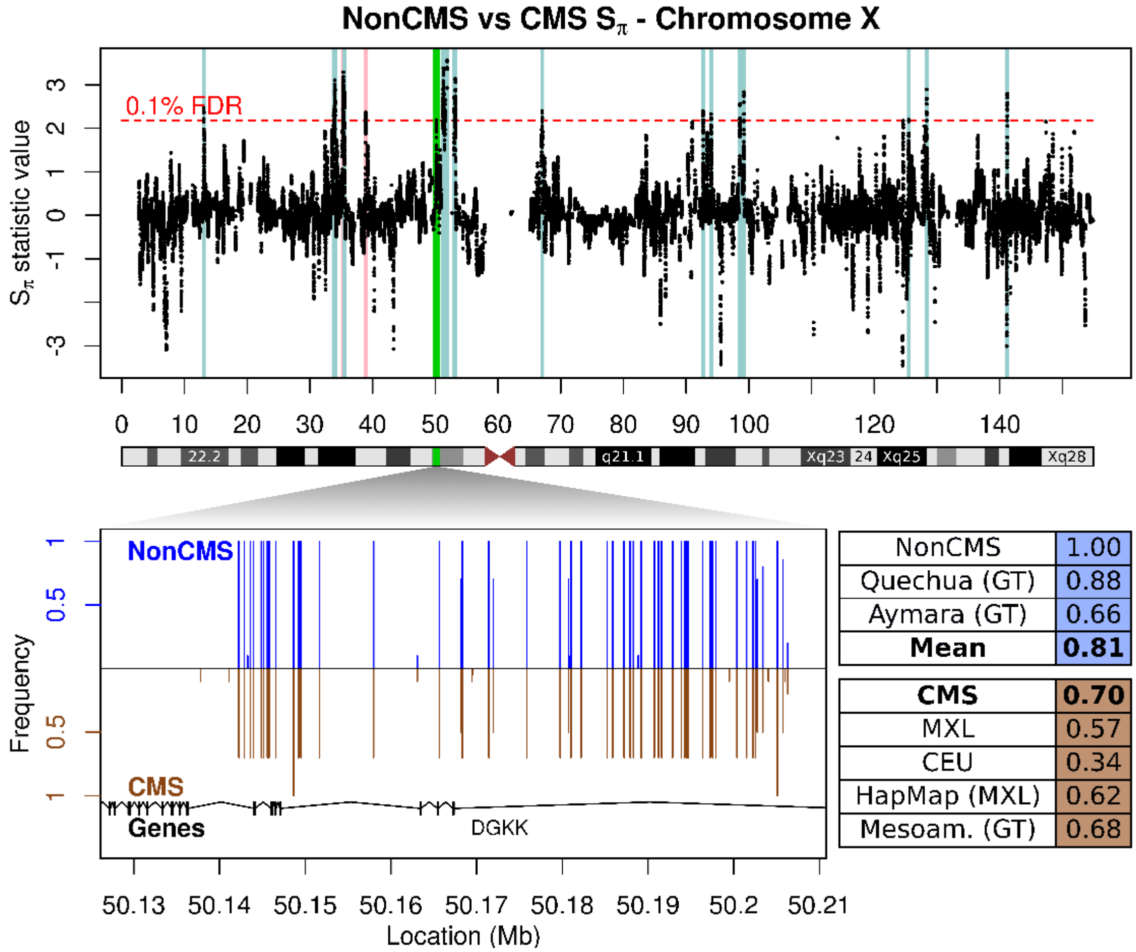
Figure S34. Profile of the first chromosome 15 (F_{ST}) candidate region. (Top) The statistic under which this region is significant ($F_{ST,MXL}$) plotted across chromosome 15. Six distinct regions exceeded the 0.1% FDR threshold, out of which two were considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 15q21.1). As can be seen, the haplotype of interest is at a much higher frequency in the non-CMS subjects than in all sampled lowlander and mal-adapted controls.



NonCMS	0.40
Quechua (GT)	N/A
Aymara (GT)	N/A
Mean	**0.40**

CMS	**0.20**
MXL	0.06
CEU	0.00
HapMap (MKK)	0.13
Mesoam. (GT)	N/A

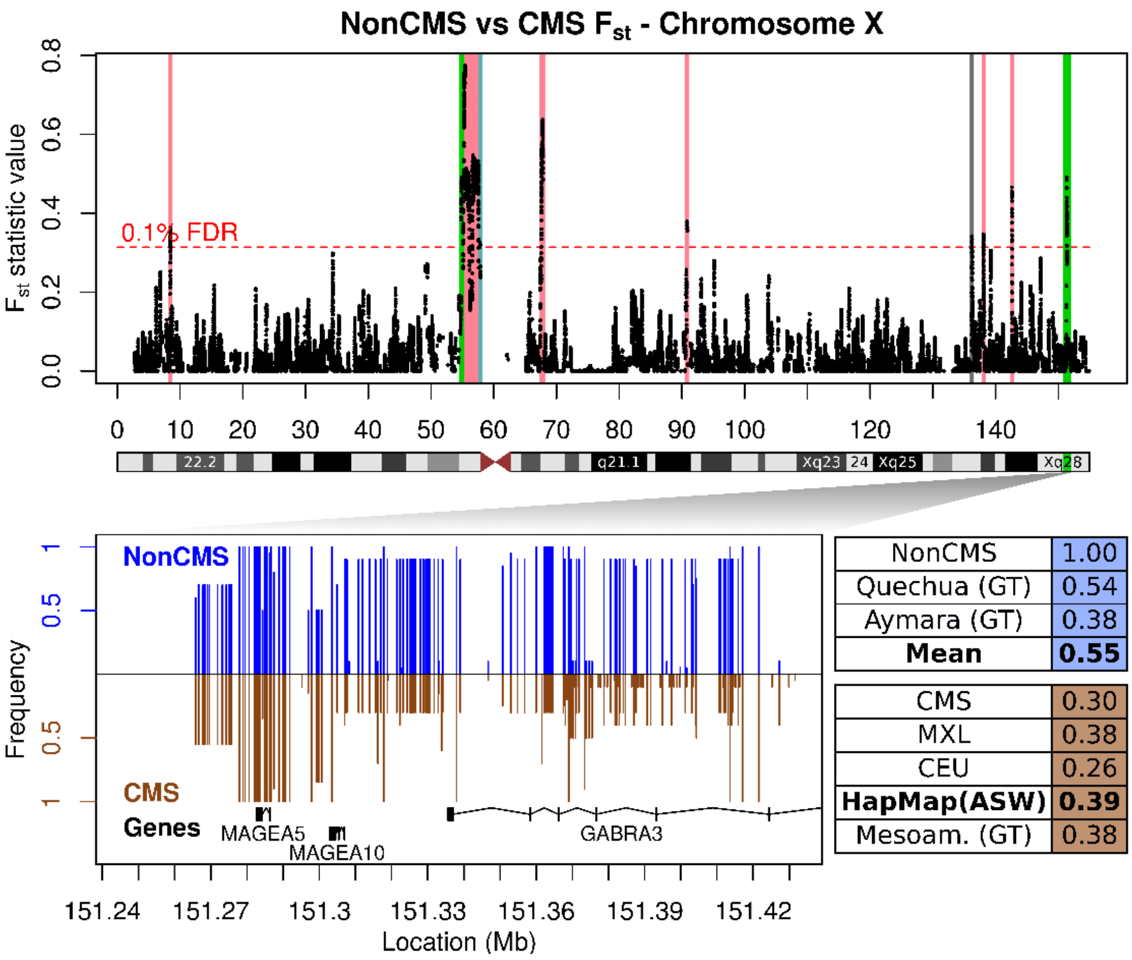
Figure S35. Profile of the chromosome 19 (*PBS*) candidate region. (Top) The statistic under which this region is significant (PBS_{MXL}) plotted across chromosome 19. One distinct region exceeded the 0.1% FDR threshold and was considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of 19p13.11). As can be seen, the haplotype in this region is at a much higher frequency in the non-CMS population relative to all sampled lowlander and mal-adapted controls.



NonCMS	1.00
Quechua (GT)	0.88
Aymara (GT)	0.66
Mean	**0.81**

CMS	**0.70**
MXL	0.57
CEU	0.34
HapMap (MXL)	0.62
Mesoam. (GT)	0.68

Figure S36. Profile of the chromosome X (S_π) candidate region. (Top) The statistic under which this region is significant ($S_{\pi,\text{CMS}}$) plotted across chromosome X. Twenty distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of Xp11.22). As can be seen, there is a completely fixed haplotype in the non-CMS population that is at a much lower frequency in all sampled lowlander and mal-adapted controls.



NonCMS	1.00
Quechua (GT)	0.54
Aymara (GT)	0.38
Mean	**0.55**

CMS	0.30
MXL	0.38
CEU	0.26
HapMap(ASW)	**0.39**
Mesoam. (GT)	0.38

Figure S37. Profile of the chromosome X (F_{ST}) candidate region. (Top) The statistic under which this region is significant ($F_{ST,CMS}$) plotted across chromosome X. Fourteen distinct regions exceeded the 0.1% FDR threshold, out of which one is considered prioritized. (Bottom) SNP frequencies in the area encompassing this region (part of Xq28). As can be seen, the haplotype is completely fixed in the non-CMS population, but at much lower frequencies in all sampled lowlander and mal-adapted controls.

Table S1. Chronic mountain sickness score of CMS and non-CMS patients.

Group	**Subject ID**	**Age**	**Dizziness**	**Physical weakness**	**Mental fatigue**	**Anorexia**	**Muscle weakness**	**Joint pain**	**Breathlessness**	**Palpitations**	**Disturbed sleep**	**Cyanosis**	**Injected conjuncative**	**Dilation**	**Paresthesia**	**Headache**	**Tinnitus**	**Hct score**	**Sat score**	**Hct**	**Sat (%)**	**TOTAL**
CMS	**76-AS**	28	1	1	1	0	0	1	2	2	2	2	0	2	2	3	3	3	3	84	80	28
	77-JN	39	1	1	1	0	0	0	2	2	2	2	2	2	2	3	0	3	0	67	84	23
	55-EC	34	0	1	1	0	0	1	2	2	0	2	2	2	0	3	0	3	3	65	84	22
	50-JB	37	0	0	0	1	0	1	2	0	0	2	2	0	2	3	0	3	3	76	80	19
	40-JI	43	0	1	1	0	0	1	2	2	2	2	2	2	0	3	3	3	3	80	71	27
	43-AM	42	1	1	1	1	1	1	2	2	2	2	2	0	0	3	0	3	0	78	83	22
	16-EJ	41	0	0	1	0	1	1	0	0	0	2	2	2	2	3	3	3	3	78	80	23
	98-AC	50	1	0	1	0	0	0	2	2	2	2	2	0	0	3	3	3	3	70	80	24
	96-JP	27	1	1	1	1	0	0	2	2	2	2	2	0	2	3	3	3	0	70	82	25
	93-BS	41	1	1	0	0	0	1	2	2	0	2	2	0	2	3	3	3	3	70	76	25
Non-CMS	**56-AM**	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	86	0
	63-PF	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	84	0
	109-EP	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	90	0
	82-JC	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	52	83	0
	53-JR	23	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	52	90	3
	52-AA	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	50	88	0
	72-DA	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	54	87	0
	103-RQ	43	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	54	85	1
	31-CE	25	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	52	91	2
	7-PCE	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	88	0

Table S2. Sequencing depth, coverage and variant call statistics per individual in the sample.

	Sample	**Reads (K)**	**% Mapped**	**Mean Depth**	**% Coverage**	**Final Variants**
non-CMS	56-AM	352,075	96.84%	18.59x	99.34%	3,019,403
non-CMS	63-PF	766,344	96.42%	40.27x	99.54%	2,989,899
non-CMS	109-EP	325,790	96.65%	17.11x	99.35%	3,005,552
non-CMS	82-JC	805,102	96.29%	39.42x	99.55%	2,994,166
non-CMS	53-JR	808,757	96.27%	42.36x	99.54%	3,046,221
non-CMS	52-AA	755,132	95.54%	37.86x	99.53%	2,988,482
non-CMS	72-DA	763,936	96.67%	40.50x	99.56%	3,014,548
non-CMS	103-RQ	340,307	96.75%	18.08x	99.37%	2,985,187
non-CMS	31-CE	719,323	96.09%	36.37x	99.51%	2,976,551
non-CMS	7-PCE	755,711	96.37%	37.62x	99.54%	3,008,044
CMS	76-AS	792,067	96.93%	40.52x	99.56%	3,030,933
CMS	77-JN	758,303	96.81%	40.46x	99.52%	3,024,571
CMS	55-EC	341,837	96.68%	17.84x	99.34%	3,016,358
CMS	50-JB	776,939	96.36%	40.46x	99.53%	2,996,461
CMS	40-JI	771,554	96.54%	36.89x	99.57%	3,051,450
CMS	43-AM	725,629	96.27%	36.86x	99.53%	3,013,376
CMS	16-EJ	353,283	96.90%	18.79x	99.38%	2,993,760
CMS	98-AC	755,715	96.17%	39.04x	99.53%	3,035,407
CMS	96-JP	787,070	96.29%	40.17x	99.55%	3,041,991
CMS	93-BS	777,315	95.97%	39.97x	99.59%	3,018,337

Table S3. UAS-RNAi lines used for hypoxia tolerance test.

Gene Symbol	Fly ortholog	Stock #.	FlyBase Genotype
SENP1	CG32110	107634	P{KK104138}VIE-260B
		34064	w[1118]; P{GD10462}v34064
		34062	w[1118]; P{GD10462}v34062
ANP32D	Mapmodulin	100283	P{KK107456}VIE-260B
		49385	w[1118]; P{GD11765}v49385

Table S4. Specific primers used in real-time qPCR.

Gene Symbol	Forward	Reverse
SENP1	AAGCCCAGACTGAAAAGGGT	CAGAGTTAAGGACCAGTGGGG
ANP32D	GAGCCTCAGTGGGCCTAGAA	GGGGCTCTATTGTGCTGAGG
PFKM	TCCCAAGGACAATCTGCAAGAA	TGCATGGTGCTCTTCATGGG
GAPDH	GCACCGTCAAGGCTGAGAAC	CGCCCCACTTGATTTTGG

Note: Specific primers were designed using Primer-BLAST at (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome)