**“IBD, source-sink population dynamics and dispersal facilitation by trade routes: impact on population genetic structure of a stored grain pest”**

Supplementary material

**Sequencing data**

**Table S1.** Illumina HiSeq2000 output summary.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Library | Samples/lane | Density (K/mm2) | Cluster PF (%) | Phas/Prephas | Read (M) | Reads PF | % >=30 | Yield (G) |
| RD\_5 | 80 | 846+63 | 88.12+8.87 | 0.128/0.051 | 233.94 | 205.52 | 92.3 | 20.6 |
| RD\_6 | 80 | 860+76 | 88.33+5.42 | 0.125/0.047 | 237.70 | 209.15 | 90.8 | 20.6 |
| RD\_7 | 96 | 815+50 | 85.42+6.44 | 0.138/0.041 | 225.19 | 191.71 | 90.3 | 19.2 |

**SNP discovery and filtering strategies**

**Table S2.** Filtering parameters used in population module of STACKS.

|  |  |
| --- | --- |
| Input parameters | Parameter description |
| -min\_maf **0.05** | Minimum minor allele frequency required to process a nucleotide site at a locus |
| --max\_obs\_het **0.5** | Maximum heterozygosity required to process a nucleotide site at a locus |
| -p **12** (100%) | Minimum number of populations a locus must be present in to process a locus |
| -r **0.85** (85%) | Minimum number of individuals in a population required to process a locus for that population |
| -m **3** | Minimum stack depth required for individuals at a locus |

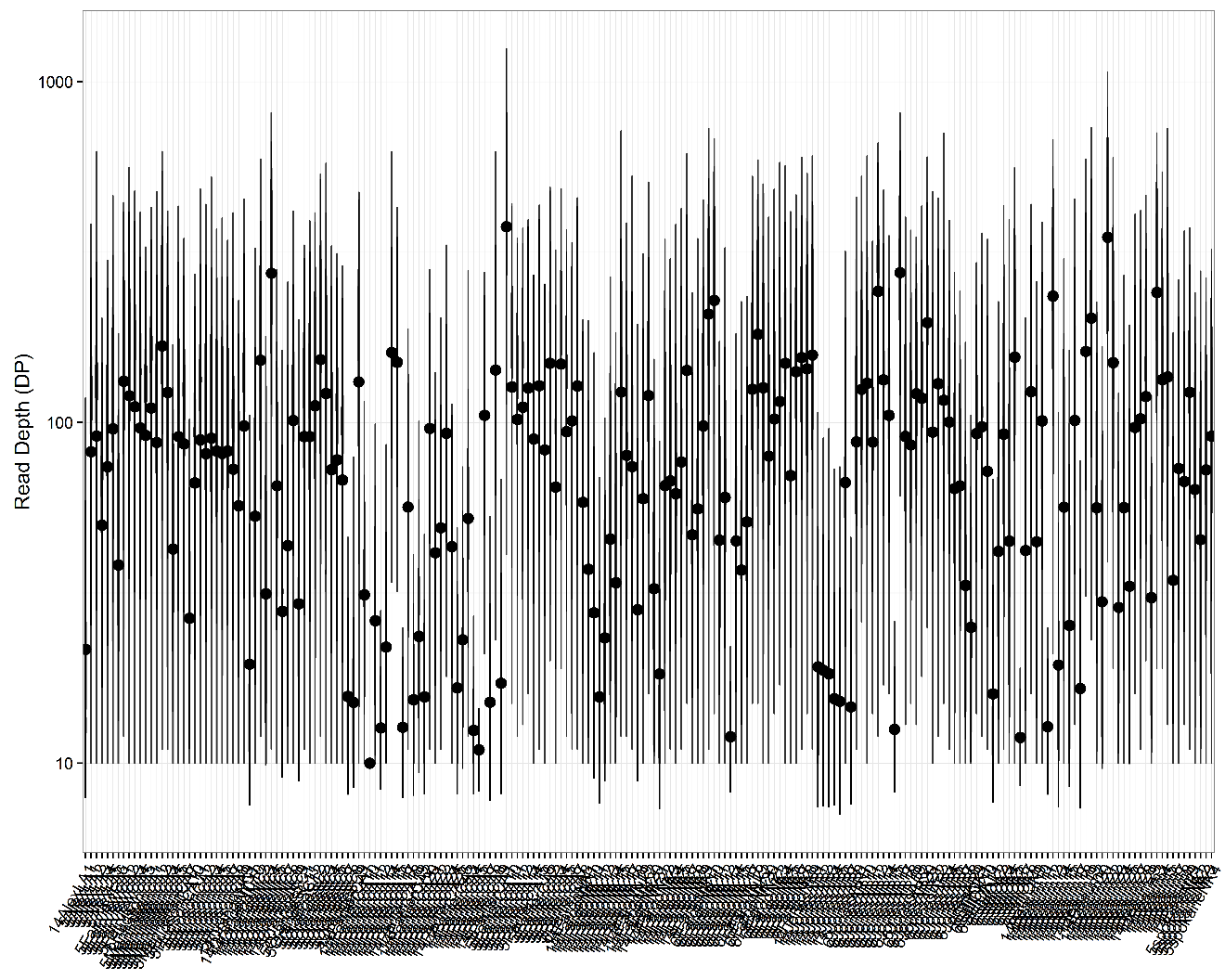


Figure S1. Depth of coverage (DP) *per* *Rhyzopertha dominica* beetle. Data refers to SNP coverage after STACKS population module filtered for minimum 10x coverage. Average coverage per individual (ID in the x axis) is 35x coverage.

**Diversity**

**Table S3.** STACKS summary statistics for each population. Variant sites are the number of sites across all population that are variant. Polymorphic sites are the number of variant sites in the current population. Parameters were calculated based on 209 beetles.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Population | Sites | Variant Sites | Polymorphic Sites | % Polymorphic *Loci* |
| LA | 95416 | 1156 | 894 | 0.9369 |
| KS1 | 95416 | 1156 | 816 | 0.8552 |
| KS2 | 95416 | 1156 | 895 | 0.938 |
| KS3 | 95416 | 1156 | 1013 | 1.0617 |
| OK | 95416 | 1156 | 1008 | 1.0564 |
| TX | 95416 | 1156 | 1073 | 1.1245 |
| AR | 95416 | 1156 | 1127 | 1.1811 |
| SC | 95416 | 1156 | 1091 | 1.1434 |
| CA | 95416 | 1156 | 1114 | 1.1675 |
| WA | 95416 | 1156 | 681 | 0.7137 |
| ND | 95416 | 1156 | 1137 | 1.1916 |

**Table S4.** STACKS summary statistics for each population. **NI** is the number of individuals sampled in this population at this site. **P** frequency of the most frequent allele; **H0** is the observed heterozygosity; **HE** is the expected heterozygosity, **π** is an estimate of nucleotide diversity, and **FIS** is the inbreeding coefficient of an individual (I) in relation to the subpopulation (S). Reported statistics were calculated taking in account all positions (variant and fixed) in each population. Parameters were calculated based on 209 beetles and 1156 variant sites (**SUP2**).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Population | NI | P | HO | HE | π | FIS |
| LA | 6.9555±0.0007 | 0.9981±0.0001 | 0.0027±0.0001 | 0.0027±0.0001 | 0.0029±0.0001 | 0.0005±0.0007 |
| KS1 | 5±0 | 0.9981±0.0001 | 0.0028±0.0001 | 0.0027±0.0001 | 0.003±0.0001 | 0.0004±0 |
| KS2 | 6.9403±0.0008 | 0.998±0.0001 | 0.0026±0.0001 | 0.0028±0.0001 | 0.003±0.0001 | 0.001±0.0008 |
| KS3 | 10.9443±0.0007 | 0.9979±0.0001 | 0.0028±0.0001 | 0.003±0.0001 | 0.0031±0.0001 | 0.0008±0.0007 |
| OK | 10.2522±0.0014 | 0.998±0.0001 | 0.0027±0.0001 | 0.0029±0.0001 | 0.003±0.0001 | 0.0011±0.0014 |
| TX | 27.4599±0.0032 | 0.998±0.0001 | 0.0026±0.0001 | 0.0028±0.0001 | 0.0029±0.0001 | 0.0008±0.0032 |
| AR | 28.6641±0.0036 | 0.998±0.0001 | 0.0026±0.0001 | 0.003±0.0001 | 0.003±0.0001 | 0.0014±0.0036 |
| SC | 20.3861±0.0025 | 0.998±0.0001 | 0.0029±0.0001 | 0.003±0.0001 | 0.003±0.0001 | 0.0006±0.0025 |
| CA | 37.1424±0.0046 | 0.9981±0.0001 | 0.0024±0.0001 | 0.0029±0.0001 | 0.0029±0.0001 | 0.0019±0.0046 |
| WA | 4±0 | 0.9981±0.0001 | 0.0027±0.0001 | 0.0025±0.0001 | 0.0029±0.0001 | 0.0003±0 |
| ND | 40.8016±0.0034 | 0.998±0.0001 | 0.0028±0.0001 | 0.003±0.0001 | 0.0031±0.0001 | 0.0011±0.0034 |

**Table S5.** STACKS summary statistics for each population. **NI** is the number of individuals sampled in this population at this site. **P** frequency of the most frequent allele; **H0** is the observed heterozygosity; **HE** is the expected heterozygosity, **π** is an estimate of nucleotide diversity, and **FIS** is the inbreeding coefficient of an individual (I) in relation to the subpopulation (S). Reported statistics were calculated taking in account only variant positions in each population. Parameters were calculated based on 209 beetles and 1156 variant sites (**SUP2**).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Population | NI | P | HO | HE | π | FIS |
| LA | 6.904±0.0087 | 0.8441±0.0041 | 0.2267±0.0057 | 0.2252±0.0049 | 0.2428±0.0053 | 0.0411±0.0087 |
| KS1 | 5±0 | 0.8413±0.0043 | 0.2334±0.0064 | 0.2234±0.0052 | 0.2483±0.0058 | 0.0308±0 |
| KS2 | 6.9083±0.0085 | 0.8358±0.0043 | 0.2171±0.0058 | 0.2327±0.0051 | 0.2509±0.0055 | 0.0823±0.0085 |
| KS3 | 10.9144±0.0082 | 0.8291±0.004 | 0.2345±0.005 | 0.2468±0.0046 | 0.2587±0.0049 | 0.0653±0.0082 |
| OK | 10.2327±0.0124 | 0.8375±0.0038 | 0.2188±0.005 | 0.2381±0.0046 | 0.2504±0.0048 | 0.0911±0.0124 |
| TX | 27.2405±0.031 | 0.8389± 0.039 | 0.2186±0.0045 | 0.2345±0.0046 | 0.239±0.0047 | 0.0621±0.031 |
| AR | 28.2128±0.0353 | 0.8348±0.0036 | 0.2177±0.0041 | 0.2461±0.0042 | 0.2506±0.0043 | 0.1141±0.0353 |
| SC | 20.0865±0.026 | 0.8327±0.0038 | 0.2353±0.0048 | 0.2452±0.0044 | 0.2515±0.0045 | 0.0536±0.026 |
| CA | 36.6384±0.0404 | 0.8411±0.0037 | 0.1954±0.0038 | 0.236±0.0044 | 0.2392±0.0044 | 0.1561±0.0404 |
| WA | 4±0 | 0.8459±0.0048 | 0.2264±0.0076 | 0.2085±0.0057 | 0.2383±0.0065 | 0.028±0 |
| ND | 40.3071±0.0358 | 0.8313±0.0036 | 0.228±0.004 | 0.2513±0.0041 | 0.2544±0.0041 | 0.0897±0.0358 |

Diversity differences, FST and effective migrant estimations

**Table S6.** Population average pairwise differences calculated in Arlequin 3.5.2.2. The above diagonal contains the average number of pairwise differences between populations (πXY). The diagonal elements are the average number of pairwise differences within population (πX) (grey color). The below diagonal contains the corrected average pairwise difference (πXY -( πX + πY)/2).

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CA | SC | ND | WA | OK | KS | TX | LA | AR |
| CA | 110.93646 | 126.12311 | 124.02439 | 135.27439 | 122.67738 | 130.96978 | 125.01093 | 128.79965 | 124.93008 |
| SC | 5.75418 | 129.80139 | 133.15172 | 146.45238 | 129.66017 | 140.33126 | 130.75780 | 134.30952 | 132.05556 |
| ND | 4.45329 | 4.14815 | 128.20575 | 143.95640 | 129.23573 | 138.48989 | 130.91259 | 136.00498 | 131.90543 |
| WA | 9.43116 | 11.17668 | 9.47852 | 140.75000 | 141.71591 | 152.56522 | 144.63362 | 149.42857 | 144.62917 |
| OK | 5.04032 | 2.59065 | 2.96403 | 9.17208 | 124.33766 | 136.41897 | 128.65987 | 131.79870 | 128.11364 |
| KS | 2.91314 | 2.84216 | 1.79861 | 9.60181 | 1.66174 | 145.17681 | 138.72414 | 143.57764 | 139.13478 |
| TX | 9.13133 | 5.44573 | 6.39834 | 13.84725 | 6.07967 | 5.72436 | 120.82275 | 134.55911 | 129.81494 |
| LA | 5.68856 | 1.76597 | 4.25925 | 11.41071 | 1.98701 | 3.34638 | 6.50488 | 135.28571 | 134.92857 |
| AR | 5.57061 | 3.26362 | 3.91131 | 10.36292 | 2.05356 | 2.65513 | 5.51233 | 3.39447 | 127.78249 |

**Table S7.** Population pairwise FST calculated in Arlequin 3.5.2.2. Distance method using pairwise FST’s for all pair of populations using 1156 SNPs and 209 beetles. All presented FST values had *p* below 0.0001.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CA | SC | ND | WA | OK | KS | TX | LA | AR |
| CA | 0 |  |  |  |  |  |  |  |  |
| SC | 0.04763 | 0 |  |  |  |  |  |  |  |
| ND | 0.03581 | 0.03129 | 0 |  |  |  |  |  |  |
| WA | 0.08946 | 0.08203 | 0.07317 | 0 |  |  |  |  |  |
| OK | 0.04426 | 0.01939 | 0.02223 | 0.07121 | 0 |  |  |  |  |
| KS | 0.02438 | 0.02009 | 0.01386 | 0.06086 | 0.01008 | 0 |  |  |  |
| TX | 0.07374 | 0.0421 | 0.04845 | 0.10829 | 0.04793 | 0.04208 | 0 |  |  |
| LA | 0.05316 | 0.01428 | 0.03348 | 0.07774 | 0.01633 | 0.0212 | 0.05293 | 0 |  |
| AR | 0.04536 | 0.0248 | 0.02964 | 0.07898 | 0.01554 | 0.01956 | 0.04243 | 0.02714 | 0 |

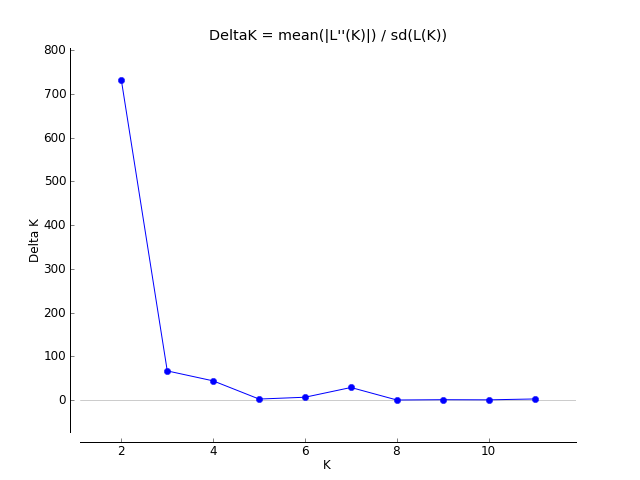
**Table S8.** Population average pairwise matrix of Slatkin linearized FST calculated in Arlequin 3.5.2.2.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CA | SC | ND | WA | OK | KS | TX | LA | AR |
| CA | 0.00000 |  |  |  |  |  |  |  |  |
| SC | 0.05001 | 0.00000 |  |  |  |  |  |  |  |
| ND | 0.03714 | 0.03230 | 0.00000 |  |  |  |  |  |  |
| WA | 0.09824 | 0.08936 | 0.07895 | 0.00000 |  |  |  |  |  |
| OK | 0.04631 | 0.01978 | 0.02274 | 0.07667 | 0.00000 |  |  |  |  |
| KS | 0.02499 | 0.02050 | 0.01406 | 0.06480 | 0.01018 | 0.00000 |  |  |  |
| TX | 0.07961 | 0.04395 | 0.05092 | 0.12144 | 0.05034 | 0.04393 | 0.00000 |  |  |
| LA | 0.05614 | 0.01448 | 0.03464 | 0.08429 | 0.01660 | 0.02166 | 0.05589 | 0.00000 |  |
| AR | 0.04752 | 0.02543 | 0.03054 | 0.08576 | 0.01579 | 0.01995 | 0.04431 | 0.02790 | 0.00000 |

**Table S9.** Population average pairwise matrix of M values (M=2Nm) calculated in Arlequin 3.5.2.2.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | CA | SC | ND | WA | OK | KS | TX | LA | AR |
| CA | \* |  |  |  |  |  |  |  |  |
| SC | 9.99861 | \* |  |  |  |  |  |  |  |
| ND | 13.46330 | 15.47941 | \* |  |  |  |  |  |  |
| WA | 5.08937 | 5.59513 | 6.33304 | \* |  |  |  |  |  |
| OK | 10.79655 | 25.28112 | 21.98758 | 6.52106 | \* |  |  |  |  |
| KS | 20.00844 | 24.38850 | 35.56377 | 7.71604 | 49.09897 | \* |  |  |  |
| TX | 6.28081 | 11.37601 | 9.81910 | 4.11718 | 9.93240 | 11.38190 | \* |  |  |
| LA | 8.90581 | 34.52448 | 14.43571 | 5.93192 | 30.12212 | 23.08486 | 8.94688 | \* |  |
| AR | 10.52287 | 19.65977 | 16.37143 | 5.83054 | 31.66639 | 25.06022 | 11.28454 | 17.92194 | \* |

**Structure and Harvester structure analysis**



**Figure S2. Putative number of populations and subpopulations.** Evanno method plot shows the most likely (highest delta K) cluster number in *R. dominica*. Secondary peaks are putative substructure.

**Table S10.** Evanno method table to evaluate the most likely (highest delta K) cluster number in *R. dominica*.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| K | Rep | Mean LnP(*K*) | Stdev | Ln’(*K*) | [Ln’’(*K*)] | Delta *K* |
| 1 | 10 | -187751.2400 | 1.4323 | - | - | - |
| 2 | 10 | -185294.2300 | 1.2084 | 2457.010000 | 883.810000 | 731.418529 |
| 3 | 10 | -183721.0300 | 9.5668 | 1573.200000 | 638.730000 | 66.764946 |
| 4 | 10 | -182786.5600 | 13.8387 | 934.470000 | 646.190000 | 46.694450 |
| 5 | 10 | -182498.2800 | 23.7988 | 288.280000 | 60.570000 | 2.545089 |
| 6 | 10 | -182270.5700 | 28.5986 | 227.710000 | 188.870000 | 6.604178 |
| 7 | 10 | -181853.9900 | 19.1334 | 416.580000 | 552.340000 | 28.867817 |
| 8 | 10 | -181989.7500 | 540.9750 | -135.760000 | 154.080000 | 0.284819 |
| 9 | 10 | -181971.4300 | 975.8394 | 18.320000 | 953.130000 | 0.976728 |
| 10 | 10 | -182906.2400 | 3634.1348 | -934.810000 | 2455.560000 | 0.675693 |
| 11 | 10 | -181385.4900 | 492.5526 | 1520.750000 | 1305.630000 | 2.650742 |
| 12 | 10 | -181170.3700 | 446.9182 | 215.120000 | - | - |



**Figure S3. Strucutre plot.** Model-based clustering of 209 beetles collected in 11 locations in the United States using 1156 SNPs markers. Analysis was performed using the program STRUCTURE, with K representing the assumed number of putative populations (HKS, FKS, and MKS are shown as KS)

**PCA**



**Figure S4. Principal component Analysis (PCA)-based k-means clustering.** Calculation are based on 1156 polymorphic single nucleotide polymorphism (SNP) markers and 209 beetles. Different colors indicate predominant crop type.

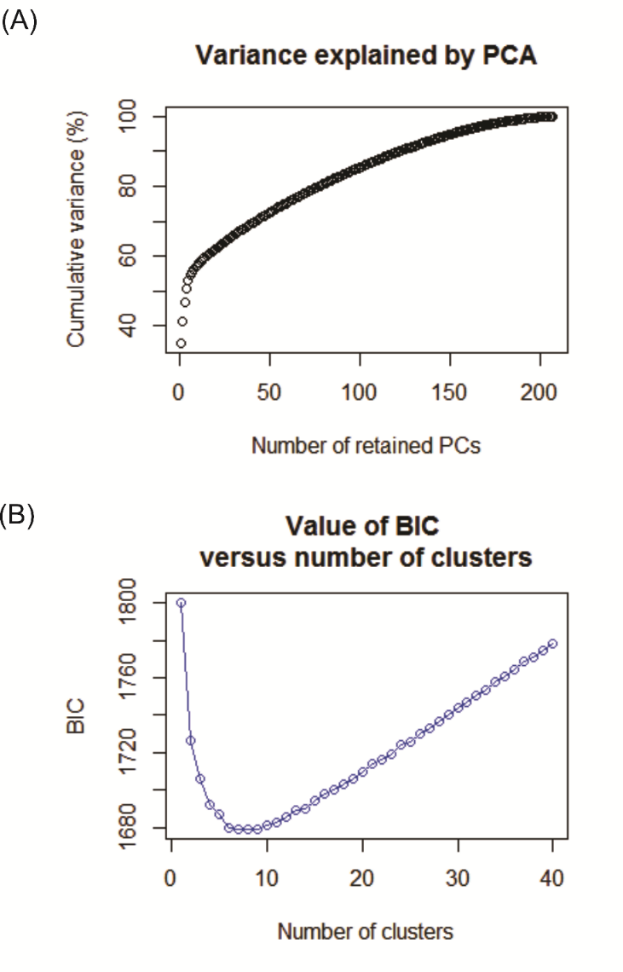


**Figure S5. Principal component Analysis (PCA)-based k-means clustering.** Calculation are based on 1156 polymorphic single nucleotide polymorphism (SNP) markers and 209 beetles. Different colors indicate the role as source or sink according to the rail transport system (Nopsa et al. 2015).

**DAPC**

To estimate the best number of clusters to explain observed variance we used Discriminant Analysis of Principal Components (DAPC) (Jombart 2008; Jombart *et al.* 2010), which optimizes the variance between groups and minimizes variance within groups (Jombart *et al.* 2010). To identify optimum number of clusters to describe the data, we used the DAPC algorithm to compare different numbers of *k* using Bayesian Information Criteria (BIC) to test for the most likely number of clusters to summarize data. First, we transformed the data using Principal Component Analysis (PCA), specifying 200 PCs to be retained; subsequently, a BIC value was calculated for each *k* value, and number of cluster determined (*k*) by picking the number of cluster associated with the BIC lowest value (k=6). Clusters were determined without sampling location information, and cluster membership probability was assessed to see if individuals collected at the same place belonged to the same cluster. Further, we analyzed the allele contribution to the PCs. Alleles can contribute differently to the observed variance and therefore might give us additional information about the relative frequency of certain alleles in different locations. All DAPC analyses were performed using *adegenet* package in R.

DAPC analysis based on allelic variation estimated k=6 as the optimal number of clusters to explain observed allelic variation. Clusters were formed without sample location information, and had high membership, with 89.5% of individuals collected at the same location belonging to the same cluster. The clusters with the highest membership, frequency of individuals collected at the same location and belonging to the same DPCA cluster, were TX, CA, ND, AR, and OK. We also analyzed the contribution of the alleles to PCA variance and the loading plot indicates that four SNPs reflect most of the variation observed among populations.



**Figure S6. Genetic diversity using multivariate DAPC analysis.** (A) The function shows cumulated variance explained by the eigenvalues of the PCA. (B) The function shows BIC values for increasing *k* value. Lower BIC value (*k*=6) indicates clusters should be retained.



Figure S7. DAPC clustering of 209 individuals from 11 locations in the United States using 5,379 SNPs markers (*K*=6). (A) The plot of the two components combination revealing of differentiation between clusters (colors represented) without sampling location information. (B) Plotting the densities of individuals on the discriminant function with different colors for DAPC clusters.

**Geographic distances**

**Table S11.** Distance matrix showing the straight-line distance (Km)—accounting for curvature of the earth—between locations calculated using R. Calculation used *GeoDistanceInMetresMatrix* function available in https://github.com/clipo/Seriation/blob/master/R/distanceCalculations.R.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | LA | TX | MKS | HKS | FKS | CA | AR | OK | ND | SC | WA |
| LA | 0 | 209 | 953 | 902 | 993 | 2561 | 529 | 684 | 1770 | 1117 | 2791 |
| TX | 209 | 0 | 1036 | 952 | 1093 | 2464 | 714 | 723 | 1879 | 1311 | 2789 |
| MKS | 953 | 1036 | 0 | 171 | 103 | 2032 | 641 | 344 | 854 | 1544 | 1923 |
| HKS | 902 | 952 | 171 | 0 | 275 | 1916 | 690 | 229 | 984 | 1621 | 1909 |
| FKS | 993 | 1093 | 103 | 275 | 0 | 2108 | 628 | 430 | 787 | 1501 | 1943 |
| CA | 2561 | 2464 | 2032 | 1916 | 2108 | 0 | 2584 | 2013 | 2196 | 3520 | 1239 |
| AR | 529 | 714 | 641 | 690 | 628 | 2584 | 0 | 575 | 1328 | 937 | 2563 |
| OK | 684 | 723 | 344 | 229 | 430 | 2013 | 575 | 0 | 1195 | 1508 | 2111 |
| ND | 1770 | 1879 | 854 | 984 | 787 | 2196 | 1328 | 1195 | 0 | 2004 | 1559 |
| SC | 1117 | 1311 | 1544 | 1621 | 1501 | 3520 | 937 | 1508 | 2004 | 0 | 3429 |
| WA | 2791 | 2789 | 1923 | 1909 | 1943 | 1239 | 2563 | 2111 | 1559 | 3429 | 0 |

**Bioclimatic variables**

**Table S12.** Nineteen bioclimatic predictions and the altitude (m) for each sample location. Variables were calculated according to Hijmans *et al.* 2005 and the maps used to extract bioclimatic information are available at the Bioclim website; (http://www.worldclim.org/bioclim).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| POP | m | WC01 | WC02 | WC03 | WC04 | WC05 | WC06 | WC07 | WC08 | WC09 | WC10 | WC11 | WC12 | WC13 | WC14 | WC15 | WC16 | WC17 | WC18 | WC19 |
| LA | 27 | 188 | 123 | 40 | 6787 | 335 | 29 | 306 | 109 | 238 | 272 | 97 | 1463 | 165 | 98 | 15 | 422 | 304 | 328 | 422 |
| TX | 9 | 198 | 115 | 39 | 6318 | 334 | 46 | 288 | 261 | 159 | 275 | 113 | 1430 | 152 | 88 | 17 | 414 | 273 | 393 | 340 |
| MKS | 356 | 120 | 131 | 31 | 10071 | 326 | -96 | 422 | 222 | -16 | 246 | -16 | 857 | 136 | 19 | 52 | 354 | 68 | 327 | 68 |
| HKS | 480 | 129 | 133 | 31 | 9890 | 339 | -80 | 419 | 229 | -3 | 254 | -3 | 744 | 110 | 15 | 48 | 297 | 62 | 277 | 62 |
| FKS | 361 | 114 | 129 | 30 | 10195 | 321 | -101 | 422 | 220 | -24 | 241 | -24 | 885 | 122 | 22 | 49 | 342 | 75 | 336 | 75 |
| CA | 99 | 169 | 152 | 44 | 6825 | 364 | 23 | 341 | 100 | 256 | 256 | 82 | 282 | 55 | 0 | 83 | 151 | 3 | 3 | 143 |
| AR | 85 | 158 | 118 | 32 | 8524 | 336 | -22 | 358 | 159 | 264 | 264 | 43 | 1216 | 130 | 78 | 19 | 379 | 241 | 241 | 299 |
| OK | 291 | 153 | 131 | 33 | 8948 | 345 | -45 | 390 | 242 | 34 | 266 | 34 | 876 | 133 | 28 | 40 | 316 | 104 | 255 | 104 |
| ND | 273 | 51 | 121 | 25 | 12450 | 283 | -199 | 482 | 202 | -119 | 202 | -119 | 531 | 87 | 11 | 57 | 235 | 42 | 235 | 42 |
| SC | 63 | 178 | 134 | 42 | 6804 | 333 | 15 | 318 | 263 | 134 | 263 | 88 | 1207 | 137 | 63 | 23 | 402 | 220 | 402 | 287 |
| WA | 676 | 84 | 118 | 34 | 7799 | 285 | -59 | 344 | -8 | 182 | 186 | -13 | 467 | 64 | 17 | 39 | 185 | 60 | 72 | 169 |

**Train rail and crops variables**

**Table S13.** State grain rail and crop production statistical summary. Shipments and receipts were compiled by the U.S. Department of Agriculture, Agriculture Marketing Service and authored by Prater et al. 2013 available at <http://dx.doi.org/10.9752/TS066.06-2013>. Data is presented as volume of grain x 106 tons. Wheat and rice acreage data were extracted from 50 km buffer using ArcGIS 10.3.1 (https://www.nass.usda.gov/Charts\_and\_Maps/Crops\_County/).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| STATE | SHIPMENTS | RECEIPTS | WHEAT | RICE | CROP |
| WA | 7.381341 | 27.564503 | 1884660.389 | 0 | Wheat |
| TX | 10.462916 | 55.590223 | 0 | 745478.6836 | Rice |
| CA | 2.2154043 | 10.590569 | 231129.5767 | 0 | Wheat |
| LA | 0.172832 | 4.431420 | 254053.2649 | 740190.1715 | Rice |
| AR | 0.356 | 0.166 | 185328.6326 | 796881.8453 | Rice |
| OK | 12.053613 | 2.637918 | 1845491.138 | 0 | Wheat |
| SC | 0 | 1.172296 | 0 | 0 | Wheat |
| KS | 35.298760 | 3.314456 | 1909107.531 | 0 | Wheat |
| ND | 47.047967 | 3.683122 | 364691.9368 | 0 | Wheat |

**Modeling parameters and variables**

**Table S14.** Modeling parameters and variables. Model selection was performed in R using *glmuli* and *leaps* packages.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model parameters | | | | Reference |
|  | | | |  |
| Response variable | M | |
| Explanatory variables | Climate (PCA scores) | WC01 "Annual Mean Temperature"  WC02 "Mean Diurnal Range"  WC03 "Isothermality"  WC04 "Temperature Seasonality"  WC05 "Max Temperature of Warmest Month"  WC06 "Min Temperature of Coldest Month"  WC07 "Temperature Annual Range"  WC08 "Mean Temperature of Wettest Quarter"  WC09 "Mean Temperature of Driest Quarter"  WC10 "Mean Temperature of Warmest Quarter"  WC11 "Mean Temperature of Coldest Quarter"  WC12 "Annual Precipitation"  WC13 "Precipitation of Wettest Month"  WC14 "Precipitation of Driest Month"  WC15 "Precipitation Seasonality"  WC16 "Precipitation of Wettest Quarter"  WC17 "Precipitation of Driest Quarter"  WC18 "Precipitation of Warmest Quarter"  WC19 "Precipitation of Coldest Quarter” | | Hijmans *et al.* 2005 |
| Agriculture/production | RICE “rice production ‘acreages’ in a 50 Km buffer”  WHEAT “wheat production ‘acreage’ in a 50 Km buffer” | | Prater *et al.* 2013 |
| Transportation | SHIPMENTS “Volume of grain ship from that location”  RECEIPTS “volume of grain received from other locations  CEN “centrality” | | Nopsa *et al.* 2015 |
| Geographical | DIST “Geographic distance” | | Hijmans *et al.* 2005 |
| level | 1 | “no interactions allowed” | |  |
| method | “h” | Exhaustive screening | |  |
| maxsize | 6 | “maximum 6 terms to be included in candidate models” | |  |
| crit | “aicc” | Corrected Akaike IC | |  |
| confsetsize | 100 | Number of model to be looked for (confidence set) | |  |
| fitfunction | “glm” |  | |  |

**Model selection, most important variables and best models for M**



**Figure S8.** AICc values for the 200 top models for the number of effective migrants (M) of *Rhyzopertha dominica*. Horizontal red line separates models with AICc less than 2 units away from the “best” model. Five models were selected according to this criterium.



**Figure S9.** Relative importance of predictors for the average M, considering all models. Vertical red line indicates the 0.8 cutoff used to separate important variables. Most important variables were the “geographical distance” and “volume of grain received”.