

1 **List of Supplementary Material**

2
3 Supplementary Methods: Description of ddRAD-Seq genotyping

4
5 Supplementary Table 1: Genetic diversity summary statistics with standard error for each
6 estimated value and isolation-by-distance Mantel r values for brown rat populations in each of
7 four urban landscapes.

8
9 Supplementary Figure 1: Evolutionary clustering for brown rats within four independent urban
10 landscapes as described by DAPC. For each city we plot the spread of genetic clusters across one
11 or more discriminant functions (top left), the BIC value for each possible value of K where lower
12 values indicate more optimal fit (bottom left), and the posterior probability of each sample's
13 assignment to a particular cluster mapped as a pie chart (right). Note that when $K=2$ only one
14 discriminant function is used, creating a density plot rather than a scatter plot, which is used
15 when $K > 2$. For NYC, $K=1$ was optimal but we show results for $K=2$ to visualize subpopulation
16 clustering.

17
18 Supplementary Figure 2: Evolutionary clustering for rats in NOL using DAPC in a $K=3$ scenario.
19 The additional cluster identifies rats from the “French Quarter” (shown in green) as
20 differentiated from other NOL rats.

21
22 Supplementary Figure 3: Spatial neighborhoods of shared genetic variation for brown rats in four
23 cities described by MEMGENE. We show results for the eigenvector that explains the largest
24 proportion of genetic variation. Black and white circles represent different established genetic
25 groups and the size of the circle represents the magnitude of difference between them (i.e., large
26 black and large white circle are the most different).

27
28 Supplementary Figure 4: Evolutionary clustering for a subset of 125 rats from all four cities,
29 analyzed together using DAPC. Samples from each city are reliably assigned to the same cluster
30 and broad global-scale relationships are identified. The third discriminant function (not shown
31 here) separates the NOL and NYC clusters.

32 **Supplementary Methods**

33

34 We used an identical ddRADSeq approach to prepare libraries for genome-wide SNP genotyping
35 of rats from all four cities. All ddRAD-Seq work except for DNA extractions were performed in
36 the corresponding author's lab at Fordham University, following established protocols [1]. In
37 brief, genomic DNA was extracted from tail or liver samples using the Qiagen DNeasy Blood &
38 Tissue Kit with a 4 μ L RNase treatment. DNA from 500-1,000 ng for each sample was
39 restriction-digested using the MluCI and SphI enzymes and cleaned using 1.5x volume of
40 Agencourt AMPure XP or similar homemade beads. Next we ligated a P1 adapter containing one
41 of 48 unique 5-nucleotide barcodes and a P2 adapter to fragment overhangs, pooled sets of 48
42 barcoded samples, and repeated bead cleaning. We then pooled 48 barcoded samples each in
43 libraries at equimolar concentrations, and selected fragments from 340-412 bp (target = 376 bp)
44 using a Pippin Prep (Sage Science, Beverly, MA). Next, we quantified DNA concentrations in
45 each pool using a Qubit fluorometer, and then amplified each for 11 cycles with Phusion High-
46 fidelity PCR reagents (New England Biolabs, Ipswich, MA). PCR primers added sequencing
47 flowcell annealing sequences and a second pool-specific indexing barcode, so each sample had a
48 unique dual barcode combination for downstream identification. Products were bead-cleaned and
49 inspected with an Agilent 2100 BioAnalyzer before paired-end 125bp sequencing on Illumina
50 HiSeq 2500 instruments at the New York Genome Center.

51 To identify SNPs from sequence data we first used the *process_radtags* script from
52 STACKS v1.35 to assign reads to individuals [2]. We then aligned reads to the Rnor 6.0
53 reference genome using BOWTIE2 under default settings [3], removing any individuals with <
54 500 MB of aligned reads. Next we ran *pstacks* (m = 3), *cstacks* (n = 2), and *sstacks* scripts from
55 STACKS on samples from each city separately to capture within-city variation. Using the
56 *populations* script, we created data sets for each city using several parameters: retained only loci
57 found in $\geq 85\%$ of samples ($r = 0.85$); retained SNPs with a minor allele frequency $\geq 5\%$
58 ($\text{min_maf} = 0.05$); removed loci with high heterozygosity to limit effect of duplication within the
59 genome ($\text{max_het} = 0.8$) and retained only a single SNP per RADtag (`--write_single_snp`).
60 Lastly, we removed any individuals with > 50% missing data; the resultant average missing data
61 after this filtering was very low (NYC: 7.2%; NOL: 5.4%; VAN: 7.1%; SAL: 6.6%).

62 To compare SNPs across different urban rat populations, we reran *cstacks* to create a
63 multi-city catalog with 125 samples randomly chosen from each city. Computational constraints

64 precluded analyzing all samples from all cities in the same catalog. We then reran *sstacks* on all
65 samples using the multicity catalog. Next, we ran *populations* using all of the above mentioned
66 parameters to retain SNPs found in all four populations ($r = 4$) and create an “among-city” SNP
67 dataset. For each genetic dataset we also calculated a matrix of the average pairwise genetic
68 dissimilarity using the program *bed2diffs_v1*, which excludes loci with missing data for each pair
69 of individuals [4]. This genetic distance metric is analagous to the proportion of shared alleles
70 (Dps), which has been shown to perform well for analyzing connectivity at small spatial scales
71 [5].

72

73

- 74 1. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012 Double Digest
75 RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and
76 Non-Model Species. *PLoS ONE* **7**, e37135. (doi:10.1371/journal.pone.0037135)
- 77 2. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013 Stacks: an analysis
78 tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. (doi:10.1111/mec.12354)
- 79 3. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat.*
80 *Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)
- 81 4. Petkova D, Novembre J, Stephens M. 2016 Visualizing spatial population structure with
82 estimated effective migration surfaces. *Nat. Genet.* **48**, 94–100. (doi:10.1038/ng.3464.)
- 83 5. Bowcock, A, Minch, E. 1994 High Resolution of human evolutionary trees with polymorphic
84 microsatellites. *Nature* **368**, 455–457. (doi: 10.1038/368455a0)

85
86

87 **Supplementary Table 1.** Genetic diversity summary statistics with standard error for each estimated value and isolation-by-distance

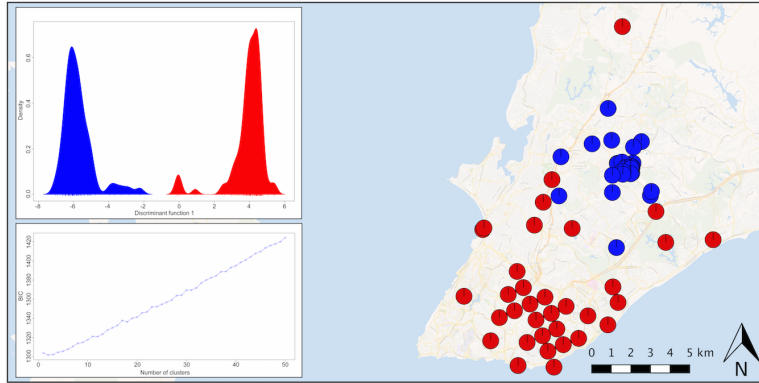
88 Mantel r values for brown rat populations in each of four urban landscapes.

89

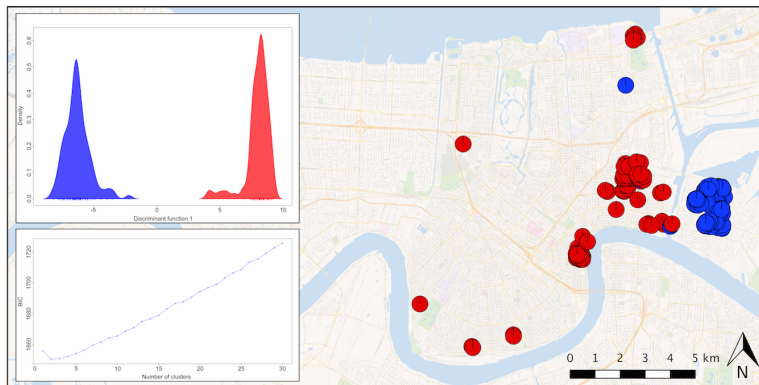
City	Samples	SNPs	Observed Heterozygosity	Expected Heterozygosity	Nucleotide Diversity (π)	Inbreeding Coefficient (F_{IS})	Isolation-by-Distance (mantel r)
New York City, USA	262	36,465	0.255 (± 0.0006)	0.307 (± 0.0007)	0.308 (± 0.0007)	0.173 (± 0.0554)	0.30 ($p = 0.001$)
Salvador, Brazil	153	27,236	0.245 (± 0.0007)	0.314 (± 0.0008)	0.316 (± 0.0008)	0.223 (± 0.0334)	0.61 ($p = 0.001$)
Vancouver, Canada	615	15,310	0.239 (± 0.0009)	0.318 (± 0.0011)	0.318 (± 0.0011)	0.243 (± 0.2099)	0.71 ($p = 0.001$)
New Orleans, USA	193	31,349	0.277 (± 0.0007)	0.314 (± 0.0007)	0.315 (± 0.0007)	0.122 (± 0.0452)	0.69 ($p = 0.001$)

90

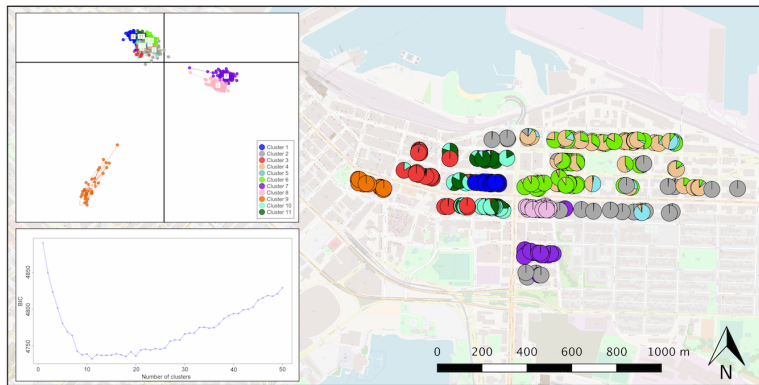
SAL



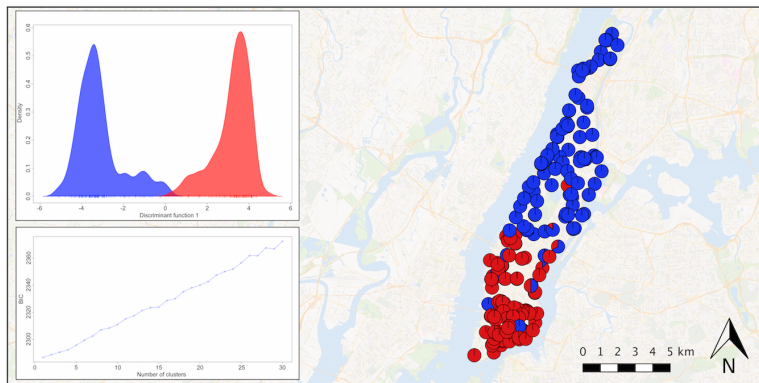
NOL



VAN



NYC

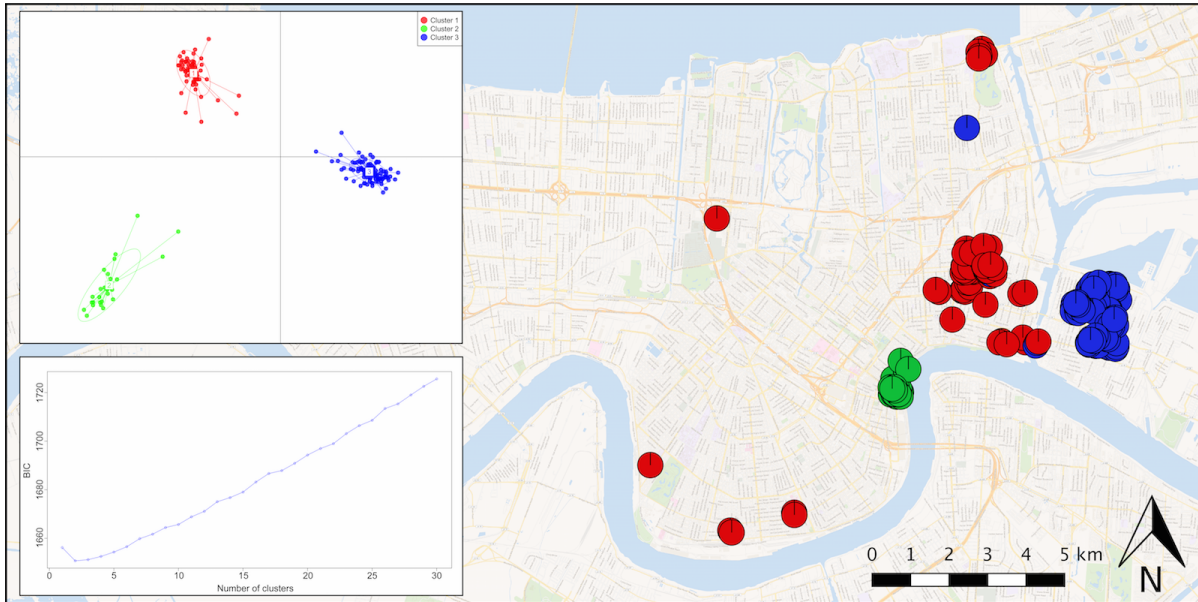


92 **Supplementary Figure 1:** Evolutionary clustering for brown rats within four independent urban
93 landscapes as described by DAPC. For each city we plot the spread of genetic clusters across one
94 or more discriminant functions (top left), the BIC value for each possible value of K where lower
95 values indicate more optimal fit (bottom left), and the posterior probability of each sample's
96 assignment to a particular cluster mapped as a pie chart (right). Note that when K=2 only one
97 discriminant function is used, creating a density plot rather than a scatter plot, which is used
98 when $K > 2$. For NYC, K=1 was optimal but we show results for K=2 to visualize subpopulation
99 clustering.

100

101

102

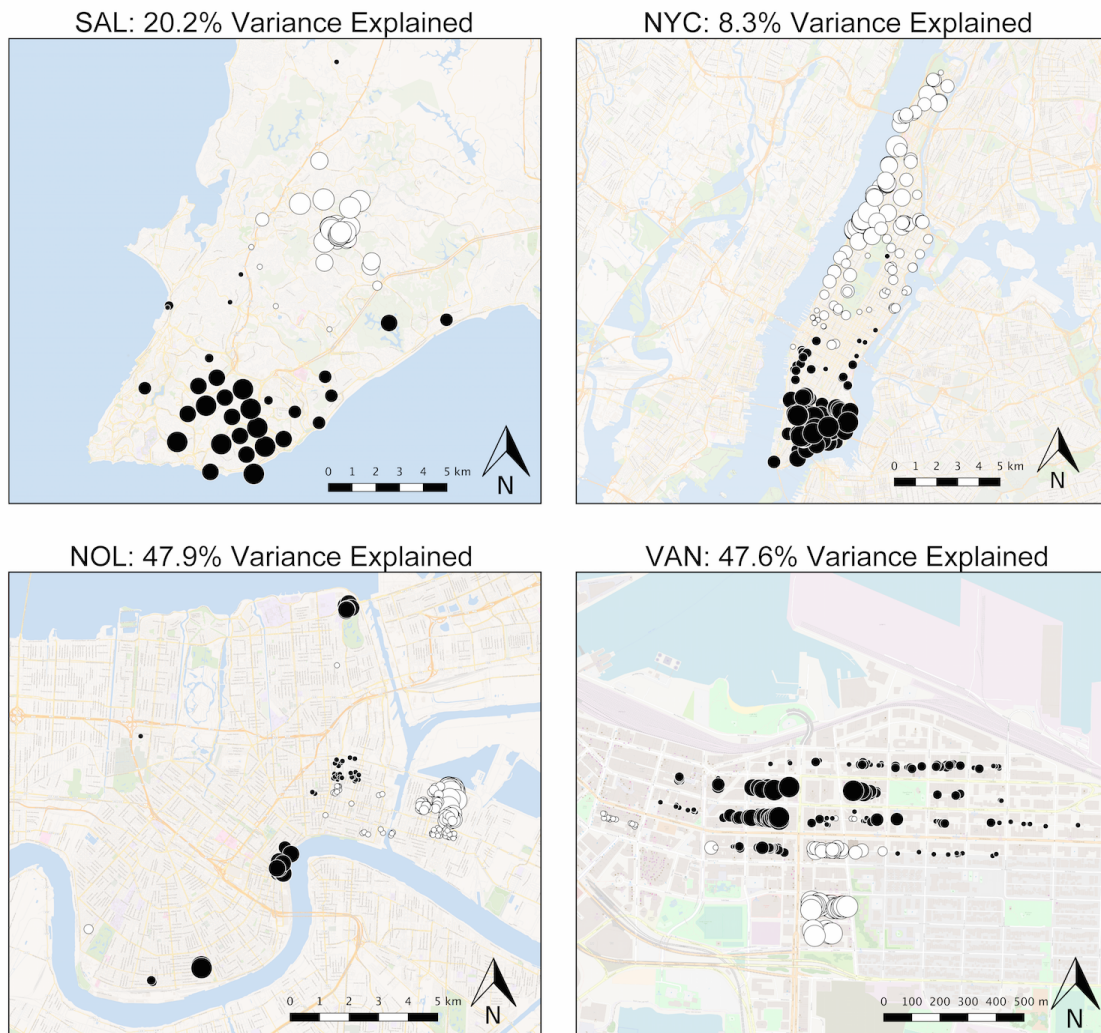


103

104

105 **Supplementary Figure 2.** Evolutionary clustering for rats in NOL using DAPC in a K=3
106 scenario. The additional cluster identifies rats from the “French Quarter” (shown in green) as
107 differentiated from other NOL rats.

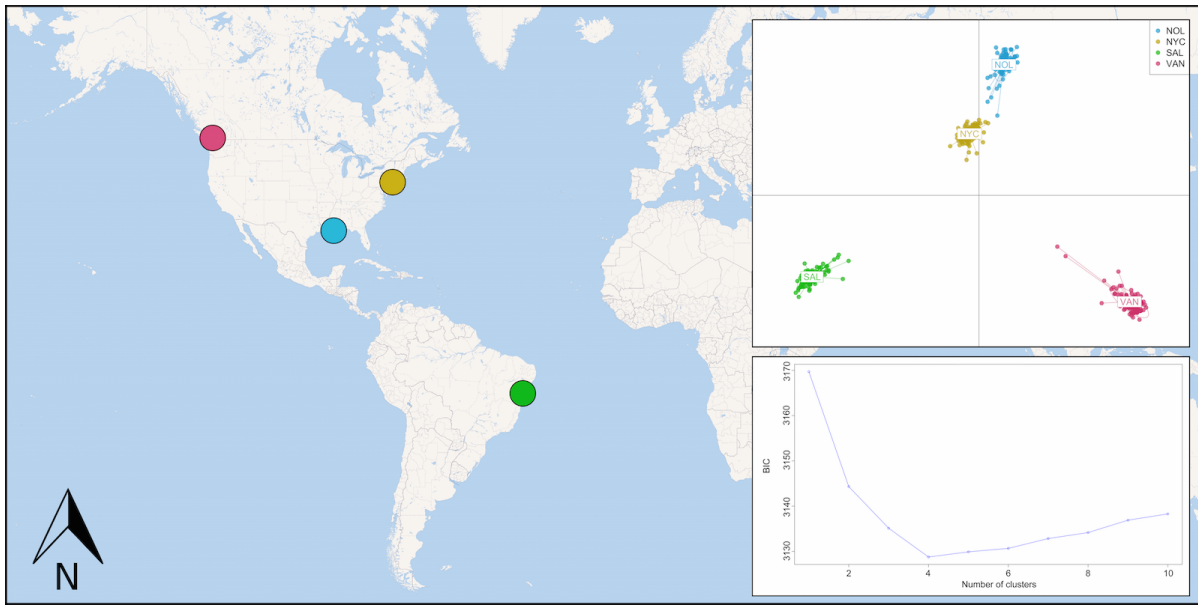
108



109

110 **Supplementary Figure 3:** Spatial neighborhoods of shared genetic variation for brown rats in
 111 four cities described by MEMGENE. We show results for the eigenvector that explains the
 112 largest proportion of genetic variation. Black and white circles represent different established
 113 genetic groups and the size of the circle represents the magnitude of difference between them
 114 (i.e., large black and large white circle are the most different).

115



116

117

118 **Supplementary Figure 4:** Evolutionary clustering for a subset of 125 rats from all four cities,
 119 analyzed together using DAPC. Samples from each city are reliably assigned to the same cluster
 120 and broad global-scale relationships are identified. The third discriminant function (not shown
 121 here) separates the NOL and NYC clusters.