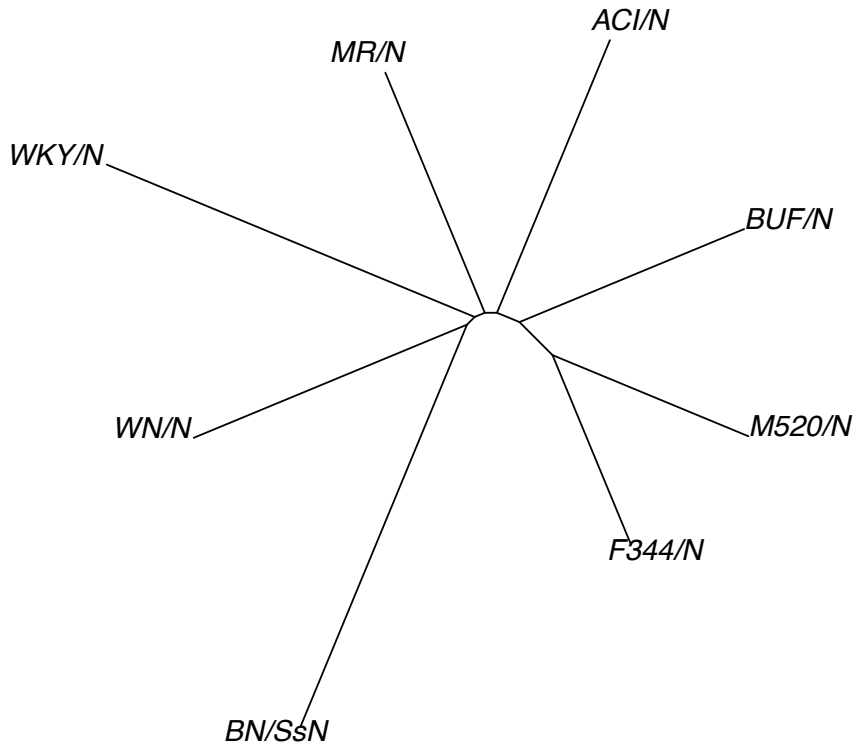
# Combined sequence-based and genetic mapping analysis of complex traits in outbred rats

**Rat Genome Sequencing and Mapping Consortium**

Amelie Baud, Roel Hermsen, Victor Guryev, Pernilla Stridh, Delyth Graham, Martin W. McBride, Tatiana Foroud, Sophie Calderari, Margarita Diez, Johan Ockinger, Amennai D. Beyeen, Alan Gillett, Nada Abdelmagid, Andre Ortlieb Guerreiro-Cacais, Maja Jagodic, Jonatan Tuncel, Ulrika Norin, Elisabeth Beattie, Ngan Huynh, William H. Miller, Anna F. Dominiczak, Daniel L. Koller, Imranul Alam, Samreen Falak, Mary Osborne-Pellegrin, Esther Martinez-Membrives, Toni Canete, Gloria Blazquez, Elia Vicens-Costa, Carme Mont-Cardona, Sira Diaz-Moran, Adolf Tobena, Oliver Hummel, Diana Zelenika, Kathrin Saar, Giannino Patone, Anja Bauerfeind, Marie-Therese Bihoreau, Matthias Heinig, Young-Ae Lee, Carola Rintisch, Herbert Schulz, David A. Wheeler, Kim C. Worley, Donna M. Muzny, Richard A. Gibbs, Mark Lathrop, Nico Lansu, Pim Toonen, Frans Paul Ruzius, Ewart de Bruijn, Heidi Hauser, Santosh S. Atanur, Tim J. Aitman, Paul Flicek, David J. Adams, Thomas Keane, Tomas Malinauskas, E. Yvonne Jones, Diana Ekman, Regina Lopez-Aumatell, Anna F Dominiczak, Martina Johannesson, Rikard Holmdahl, Tomas Olsson, Dominique Gauguier, Norbert Hubner *, Alberto Fernandez-Teruel*, Edwin Cuppen*, Richard Mott*, Jonathan Flint*

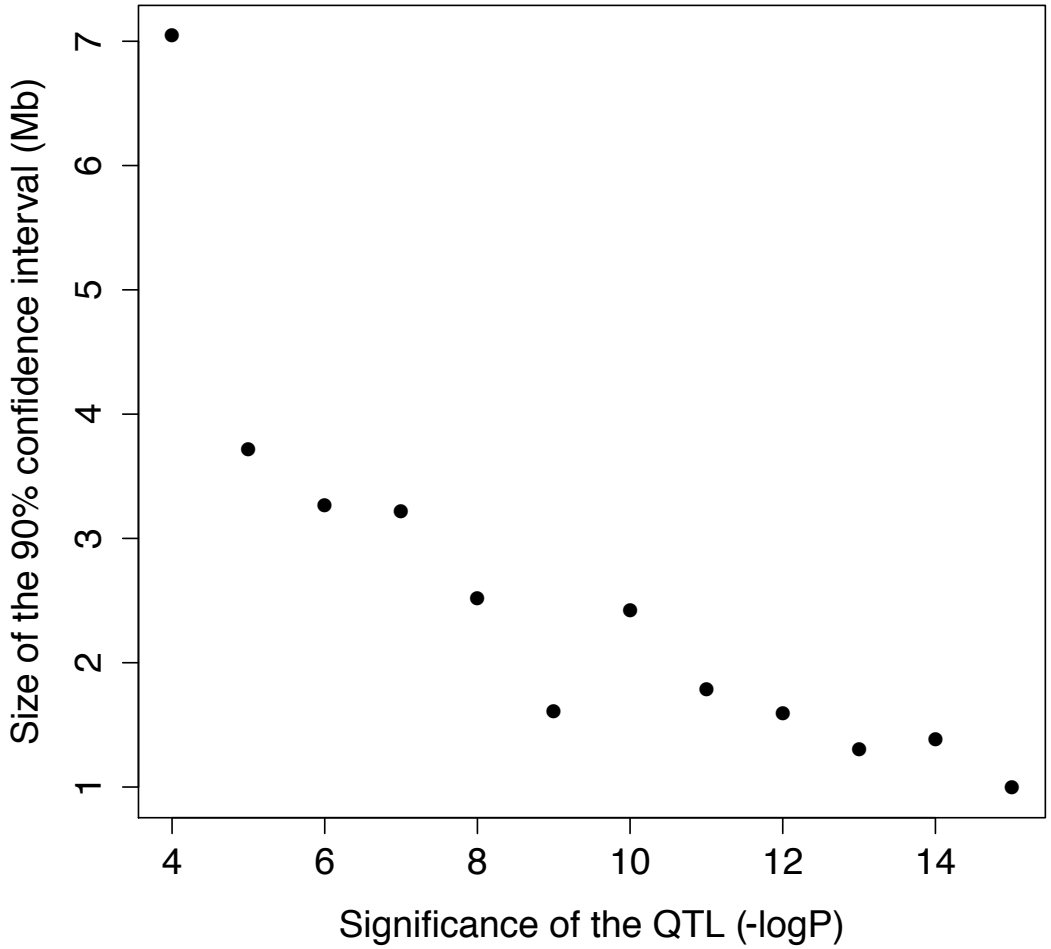
Nature Genetics: doi:10.1038/ng.2644

**Supplementary Figure 1 Pairwise distances between the rat HS founders**.
Neighbor-joining tree based on pairwise sequence distances (SNPs/kb of accessible sequence).



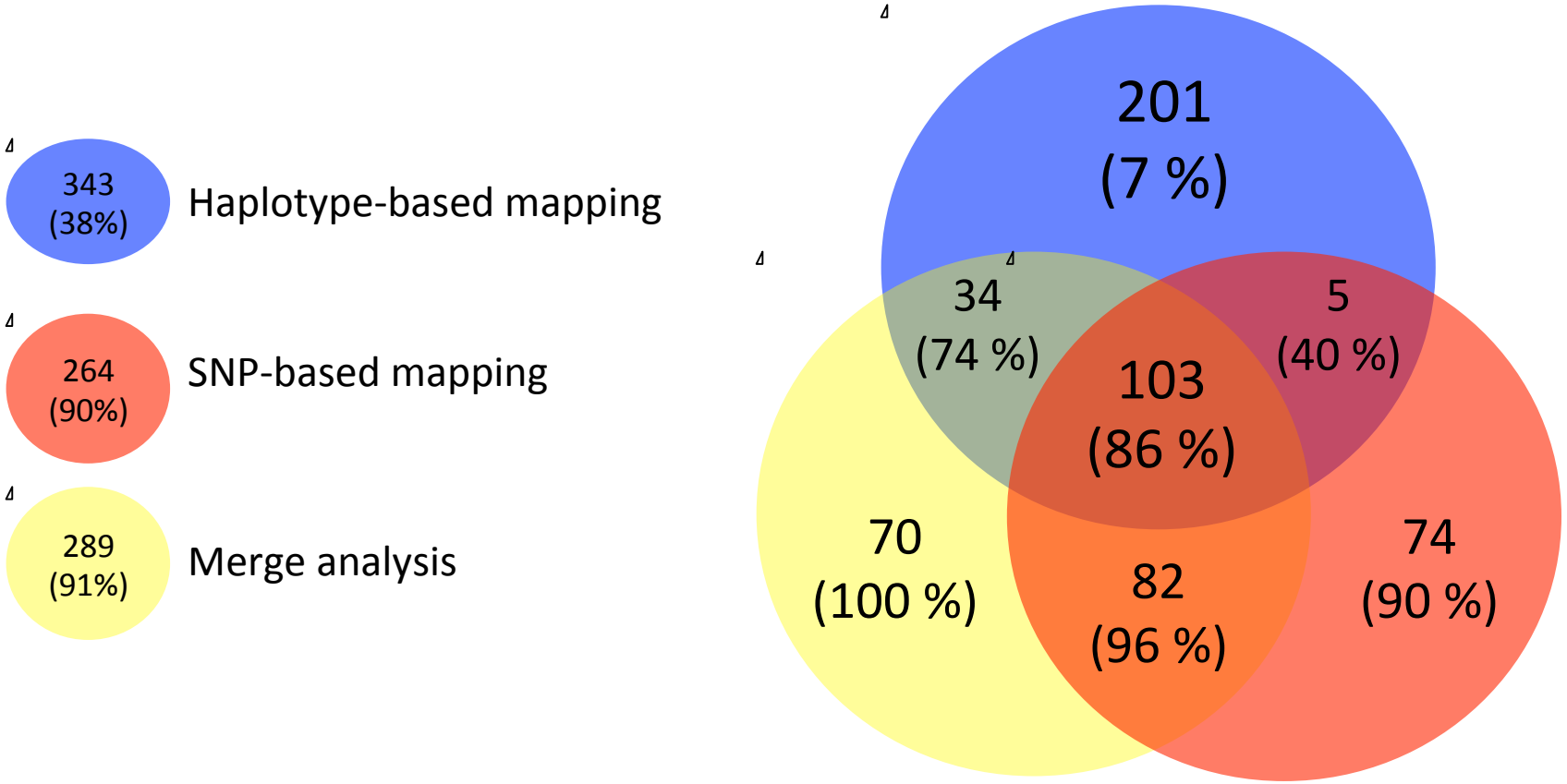

Nature Genetics: doi:10.1038/ng.2644

**Supplementary Figure 2 Size of the 90% confidence intervals as a function of the significance of the QTL.** The figure shows the size of the 90% confidence interval for the position of the QTL (most highly associated interval) as a function of its significance (negative logarithm of the p-value) based on simulations.



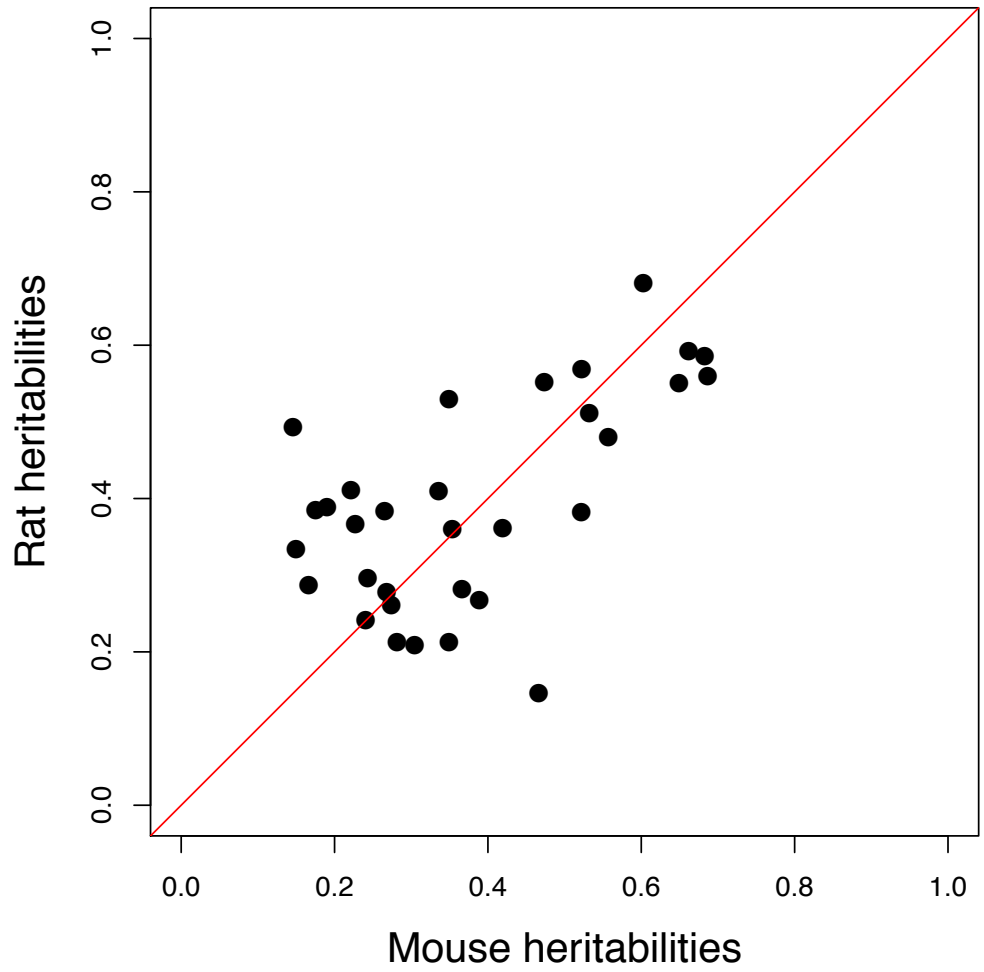

Nature Genetics: doi:10.1038/ng.2644

**Supplementary Figure 3 Concordance between the haplotype-based and SNP-based mapping methods, and the merge analysis.** The number of QTLs detected by each method is indicated, as well as the proportion of QTLs with candidate variants (in brackets). The color code and the figures for the full sets are indicated on the left side.




Nature Genetics: doi:10.1038/ng.2644

**Supplementary Figure 4 Correlation between heritability in the rat and mouse HS.** Each dot represents a measure. The red line shows the line y = x.




Nature Genetics: doi:10.1038/ng.2644

**Supplementary Table 2**
**Regions of very high interchromosomal linkage disequilibrium (mis-assemblies)**

| Chromosome | Approximate start (bp) | Approximate stop (bp) |
|---|---|---|
| 1 | 35,951,205 | 37,213,124 |
| 17 | 2,202,903 | 2,887,911 |
| 4 | 46,491,865 | 46,491,865 |
| 9 | 204,359 | 2,599,275 |
| 4 | 12,917,785 | 13,429,152 |
| 14 | 46,598,211 | 67,803,639 |
| 12 | 41,030,616 | 46,318,674 |
| 19 | 14,393,582 | 14,453,551 |

**Supplementary Table 4: Overlap between QTLs mapped in the NIH-HS and QTLs reported in the Rat Genome Database (RGD)**

| Measure | Chr. | Start HS QTL (bp) | Stop HS QTL (bp) | QTL ID in RGD | Start RGD QTL (bp) | Stop RGD QTL (bp) | Type of cross | p-value of overlap |
|---|---|---|---|---|---|---|---|---|
| Score of lesions in abdominal aorta | 5 | 164,908 | 4,164,908 | 631551 | 1,368,223 | 5,112,800 | F2 | <1/1000 |
| Total cholesterol | 1 | 65,844,068 | 78,635,862 | 631512 | 71,659,707 | 90,282,193 | | |
| Total cholesterol | 1 | 264,698,503 | 267,865,338 | 631690 | 243,502,887 | 267,910,886 | | 0.007 |
| Total cholesterol | 1 | 264,698,503 | 267,865,338 | 631835 | 226,083,572 | 267,910,886 | backcross | |
| Total cholesterol | 1 | 264,698,503 | 267,865,338 | 631836 | 244,611,037 | 267,910,886 | backcross | |
| Heart weight | 1 | 202,149,330 | 206,633,814 | 1358292 | 201,920,676 | 246,920,676 | | |
| Heart weight | 10 | 37,243,161 | 47,929,176 | 631532 | 18,217,625 | 53,793,117 | intercross | 0.034 |
| Heart weight | 13 | 56,470,137 | 67,956,785 | 1558644 | 25,198,204 | 70,198,204 | | |
| Femur midshaft polar moment of inertia | 1 | 201,265,878 | 205,189,952 | 2293654 | 176,612,333 | 221,612,333 | intercross | |
| Femur midshaft polar moment of inertia | 10 | 90,877,745 | 95,483,369 | 2293663 | 74,167,134 | 110,718,848 | intercross | 0.058 |
| Femur midshaft polar moment of inertia | 4 | 96,051,838 | 106,296,290 | 1578658 | 61,483,655 | 106,483,655 | | |
| Lumbar mineral density | 1 | 195,565,176 | 199,903,482 | 2300174 | 176,612,333 | 221,612,333 | intercross | 0.157 |
| Body weight at day immunization | 12 | 7,483,933 | 12,147,931 | 2303568 | 1 | 25,457,135 | intercross | |
| Body weight at day immunization | 2 | 129,257,158 | 133,732,902 | 1358887 | 24,474,676 | 163,154,227 | | |
| Body weight at day immunization | 2 | 138,471,008 | 142,614,192 | 1358887 | 24,474,676 | 163,154,227 | | |
| Body weight at day immunization | 2 | 216,921,362 | 220,119,428 | 1358900 | 163,154,358 | 227,150,051 | | |
| Body weight at day immunization | 2 | 129,257,158 | 133,732,902 | 1358908 | 24,474,676 | 163,154,227 | | |
| Body weight at day immunization | 2 | 138,471,008 | 142,614,192 | 1358908 | 24,474,676 | 163,154,227 | | |
| Body weight at day immunization | 3 | 23,637,280 | 36,416,292 | 1354589 | 30,253,942 | 76,620,970 | intercross | |
| Body weight at day immunization | 3 | 23,637,280 | 36,416,292 | 1354604 | 30,253,942 | 103,304,908 | intercross | |
| Body weight at day immunization | 3 | 23,637,280 | 36,416,292 | 1558654 | 6,373,335 | 26,674,263 | | 0.3 |
| Body weight at day immunization | 4 | 103,453,924 | 108,048,784 | 1357342 | 75,732,943 | 119,369,308 | | |
| Body weight at day immunization | 4 | 103,453,924 | 108,048,784 | 1549843 | 60,262,965 | 104,415,981 | | |
| Body weight at day immunization | 4 | 27,024,851 | 31,510,205 | 2303585 | 11,706,134 | 56,706,134 | intercross | |
| Body weight at day immunization | 4 | 103,453,924 | 108,048,784 | 70167 | 75,732,943 | 119,369,308 | | |
| Body weight at day immunization | 8 | 81,857,162 | 86,244,836 | 1331837 | 85,328,126 | 103,665,018 | | |
| Body weight at day immunization | 8 | 81,857,162 | 86,244,836 | 1358912 | 54,364,071 | 112,242,906 | | |
| Body weight at day immunization | 8 | 81,857,162 | 86,244,836 | 1582243 | 57,308,572 | 89,558,994 | intercross | |
| Body weight at day immunization | 4 | 103,453,924 | 108,048,784 | 1549839 | 60,262,965 | 116,780,394 | | |
| Weight loss compared to day 9 | 18 | 28,362,808 | 32,641,632 | 70178 | 13,239,641 | 58,239,641 | F2 | 0.501 |
| Weight loss compared to day 9 | 4 | 62,020,592 | 66,515,872 | 2317577 | 65,821,936 | 71,729,738 | | |
| Femur midshaft total area | 10 | 91,151,727 | 95,209,387 | 2293646 | 74,167,134 | 110,718,848 | intercross | 0.7 |

# Supplementary information for

## Combined sequence-based and genetic mapping analysis of complex traits in outbred rats

## Table of Contents


Phenotyping ..... 2
- Elevated Zero maze ..... 2
- Automated novel-cage "open field-like" activity ..... 2
- Context-conditioned fear and two-way active avoidance acquisition in the shuttlebox ..... 2
- Intraperitoneal Glucose Tolerance Test (IPGTT) ..... 2
- Blood pressure ..... 3
- Haematology ..... 3
- Basal immunology ..... 3
- Induction and Clinical Evaluation of Experimental Autoimmune Encephalomyelitis (EAE) ..... 4
- Tissue dissection ..... 4
- Serum biochemistry ..... 4
- Internal elastic lamina ruptures ..... 5
- Bone Phenotyping ..... 5
- Wound healing ..... 5

Development of a genotyping microarray ..... 6
- The pilot SNP array ..... 6
- SNP calling ..... 8
- Final Array Design ..... 9
- Genotype calling ..... 10
- Selection of a set of 500K good quality SNPs available on the RATDIV array ... 11
- Validation of Genotype Calls (on 500K SNPs) ..... 12

Selection of SNPs for this study ..... 14

Comparison between mixed models and resample model averaging ..... 15
- Scope of the methods ..... 15
- Simulation study to compare the performance of the methods ..... 15
- Comparison of the performance of the methods ..... 15

QTL calling ..... 16
- Inclusion probability thresholds for the resample model averaging analysis ... 16
- Significance thresholds for the mixed model analysis ..... 16
- Significance thresholds for the merge analysis ..... 17
- Confidence intervals ..... 17
- QTL calling ..... 17

Overlap between sets of QTLs ..... 18
- Overlap between QTLs mapped in the rat HS and QTLs catalogued in RGD ..... 18
- Overlap between the QTLs mapped in the mouse and rat HS ..... 18

Guidelines to explore the genome scans and integrated sequence data ..... 19




Nature Genetics: doi:10.1038/ng.2644

# Phenotyping

## Elevated Zero maze

The maze comprised an annular platform (105 cm diameter; 10 cm width) made of black plywood and elevated to 65 cm above the ground level. It had two open sections (quadrants) and two enclosed ones (with walls 40 cm height). The subject was placed in an enclosed section facing the wall. The apparatus was situated in a black testing room, dimly illuminated with red fluorescent light, and the behaviour was videotaped and measured outside the testing room [1,2].

## Automated novel-cage "open field-like" activity

The apparatus (Panlab, Barcelona, Spain) consisted of a horizontal surface (50 x 50 cm) provided with photo beams that detect movement and measure it automatically, loading the data in a computer. The subjects were placed in transparent Plexiglas cages (40x40x40 cm). They were situated in a white fluorescent (60 w) illuminated chamber [1,2].

## Context-conditioned fear and two-way active avoidance acquisition in the shuttlebox

Two-way active avoidance sessions were performed in three identical shuttleboxes (Letica Instr., Barcelona), each one placed in independent sound-attenuating boxes, consisting of two equal sized compartments (25 x 25 cm, 28 cm) connected by an opening (8 x 10 cm). Rats were allowed a 4-min period of familiarization to the box. Immediately after that period, a 40-trial session/rat was administered, each trial consisting of a 10-s CS (conditioned stimulus; 2400 Hz, 63-dB tone plus a 7-W small light) followed after termination by a 20-s US (unconditioned stimulus; scrambled 0.7 mA footshock) delivered through the grid floor. Crossings to the other compartment during the CS (avoidances) or US (escapes) switched off the stimuli and were followed by a 60-s inter-trial interval. Context-conditioned freezing (i.e. classically-conditioned fear) was measured by two trained observers as the time a rat spent completely motionless except for breathing movements [1,2].

## Intraperitoneal Glucose Tolerance Test (IPGTT)

Conscious rats in the post absorptive state were injected intraperitoneally with a solution of glucose (2g / kg body weight.). Blood samples were collected for glucose reading before glucose injection and 30, 60 and 120 minutes afterwards by tail tipping. Blood glucose concentration was determined using a glucose meter (Accucheck, Roche Diagnostics, Welwyn Garden City, UK). Cumulative glycemia (AUC_G) was calculated as the increment of the values of plasma glucose during the IPGTT. Incremental plasma glucose values above baseline



Nature Genetics: doi:10.1038/ng.2644

integrated over 120 min, after an injection of glucose, were used to calculate the index of glucose tolerance (DeltaG).

## Blood pressure

Systolic blood pressure was measured by tail plethysmography in conscious, restrained animals as previously described [3,4]. Briefly, rats were pre-warmed for 15 min at 30°C for tail artery vasodilation. Rats were then wrapped in a cloth for restraint and an inflatable cuff placed on their tail along with a piezoceramic transducer (Hartmann & Braun type 2). Pulse detection was visualised as a function of pressure and displayed using Microsoft Windows compatible software. An average of 6-8 pressure readings were taken for each rat per sitting.

## Haematology

Blood (350µl/sample) was collected from the lateral tail vein of anesthetized rats into a tube pretreated with the anticoagulant EDTA-2K (Sangüesa), immediately after the rats were injected with emulsified myelin oligodendrocyte glycoprotein (MOG, used to induce Experimental Autoimmune Encephalomyelitis), and stored cold until used. Full blood count measures were acquired by an automatic hemocytometer (ADVIA 120 Hematology analyzer from Bayer, Siemens Diagnostics).

## Basal immunology

Blood was collected from the lateral tail vein of anesthetized rats immediately after they were injected with emulsified MOG, into a heparinized tube (500µl/sample) and stored cold until used. A total of 20µl blood was added per well in duplicate to 96 well v-bottom polypropylene plates (BD Falcon). Staining with fluorescently labeled monoclonal antibodies (MAb) was performed on non-lysed whole blood for 20 min in the cold. MAbs were diluted in EDTA-FACS buffer (calcium- and magnesium-free PBS-D supplemented with 1% FCS, 10mM EDTA and 0.01% sodium azide) to predetermined optimal concentrations. The MAbs used were purchased from BD Pharmingen (San Diego, CA): CD45 (OX-1 FITC), RT1-B (OX-6 PE), CD45RA (OX-33 Pe-Cy5), abTCR (R73 biotin and PerCP/APC), CD8a (OX-8 FITC), CD25 (OX-39 PE/APC), CD4 (OX-35 APC/Pe-Cy5), CD28 (JJ319 FITC), RT1-A (OX-18 PE), pan-granulocyte (HIS-48 biotin/Fitc). Streptavidine (SA) -APC was used as a secondary reagent. The stained blood was incubated and washed twice at RT for 10 min in a hypertonic potassium buffer containing 8.26g ammonium chloride and 0.037g EDTA per dm3/ RBC lysis buffer, immediately after incubation with streptavidin. Washed cells were resuspended and incubated for 20 min at RT in a 2% phosphate-buffered formaldehyde solution and thereafter washed twice in EDTA-FACS buffer. Acquisition was performed on a four-color BD FACS Calibur for the first four batches and thereafter on a BD SORP LSRII Analytic Flow Cytometer, no later than 6 days post-staining. The data was analyzed with FlowJo (Tree Star Inc., Ashland, OR).


Nature Genetics: doi:10.1038/ng.2644

## Induction and Clinical Evaluation of Experimental Autoimmune Encephalomyelitis (EAE)

EAE is a highly reproducible model of multiple sclerosis with a robust clinical score scale [5,6]. Recombinant rat myelin oligodendrocyte glycoprotein (MOG), amino acids 1–125 from the N terminus, was expressed in E. coli and purified to homogeneity by metal chelate affinity chromatography [7] and ion exchange chromatography. MOG for the entire HS cohort was produced in 3 batches. Rats were anesthetized with isoflurane (Servicios Genéticos Porcinos) and immunized subcutaneously in the dorsal tail base with 200 µL inoculum containing MOG (females 50-65 µg and males 120-155 µg) in phosphate buffered-saline (PBS; Life Technologies) emulsified 1:1 with Freund's adjuvant (Sigma-Aldrich) containing 200 µg Mycobacterium tuberculosis (H37 RA, Sigma-Aldrich). Signs of EAE and body weight were monitored daily from day 9 until day 28 post-immunization (p.i.). The scale for EAE scoring was as follows: 0 = healthy; 1 = tail weakness or tail paralysis; 2 = hind leg paresis or hemiparesis; 3 = hind leg paralysis or hemiparalysis; 4 = tetraplegy, urinary, and/or fecal incontinence; and 5 = death. If severe disease (score 4) was observed for two consecutive days, the rats were sacrificed for ethical reasons. Duration of EAE was defined as the number of days with signs of disease including days after rats died/were sacrificed. Cumulative scores were defined as the sum of all scores received during the experiment including days after rats died/were sacrificed. Weight loss is a sub phenotype of EAE [8] which may depend on inflammation-mediated effects or the inability to hydrate and eat properly. WL reflects sub-clinical disease and is a quantitative trait considered to correlate well with EAE disease course. WL0 was defined as (weight at day 0 p.i. - minimum weight during the experiment)/weight at day 0 p.i.) and WL9 was defined as (weight at day 9 p.i. - minimum weight during the experiment)/weight at day 9 p.i.).

## Tissue dissection

Rats were euthanized by exsanguination under isofluorane anesthesia. Blood was obtained from cardiac puncture, then the heart was dissected out and weighed. Thereafter the ears, abdominal aorta, liver, and bones were dissected in parallel whenever possible. Other tissues were collected for future studies (thymus, brain, pituitary, spinal cord, spleen, kidneys, adrenal glands, tail). Blood was kept at room temperature for 4 hours, then at 4°C for 4 hours until it was centrifuged, the heart and liver were snap frozen in a tube in liquid nitrogen, the ears and abdominal aorta were immersed in buffered formalin, and the bones kept on ice until stored at -20°C.

## Serum biochemistry

Blood was centrifuged at 2000rpm for 20 minutes at 4°C, and the serum was separated, aliquoted and stored at -80°C. Serum was analysed on a AU400 Analyser.



Nature Genetics: doi:10.1038/ng.2644

## Internal elastic lamina ruptures

Abdominal aorta (AA) and the proximal 1 cm of the left common iliac artery (IA) were dissected out, rapidly rinsed in saline and fixed by immersion in buffered formalin. En face preparations of the unperfused AA and the attached left IA were then made. Under a dissecting microscope, arteries were cleaned, opened longitudinally and pinned out, luminal surface uppermost. The luminal surface was stained with orcein and hematoxylin to show the internal elastic lamina (IEL) and the nuclei respectively. After staining, arteries were dehydrated, unpinned, cleared and mounted on slides for microscopic observation. With this technique, as previously described [9], ruptures appear as dark grey transverse bands due to absence of the internal elastic lamina, which stains pink, and the intense staining of underlying smooth muscle cell nuclei, which are not stained in areas where the IEL is present. Ruptures were then quantified at a final magnification of x 40. For each individual, the total number of IEL ruptures in the AA and IA were recorded and each rupture was graded on a semi-quantitative scale according to its size in the circumferential direction, using a grid in the eyepiece. A final score was calculated taking into account the size of the ruptures. Thus the degree of IEL rupture in each AA and IA was expressed as a number of ruptures per artery, or as a score indicating the severity of the phenomenon. For each rat, global values were also obtained by adding those for AA and IA.

## Bone Phenotyping

For bone density and structure, femurs were placed in plastic tubes filled with 70% ethyl alcohol and centered in the gantry of a Norland Stratec XCT Research SA+pQCT (Stratec Electronics, Pforzheim, Germany). Slice measurements of 0.26 mm thickness and a voxel size of 0.07 mm were taken at the midshaft, distal femur and perpendicularly through the femoral neck. For each slice, the X-ray source was rotated through 180º of projection. Volumetric BMD (vBMD; mg/cm3), cross sectional area (CSA; mm2) and polar moment of inertia (Ip; mm4) were measured from the pQCT images. Density thresholds of 500 and 900 were used to identify mineralized bone. The femur BMC (g) was measured using DXA (PIXImus mouse densitometer, Lunar Corp., WI, USA). The femoral neck width (NW; mm) was measured in the anterior-posterior direction using digital calipers. For bone biomechanics, femurs were tested in three-point bending by positioning them on the lower supports of a three-point bending fixture and applying load at the midpoint using a material testing machine (Alliance RT/5, MTS Systems Corp., Eden Prairie, USA). Force and displacement measurements were collected every 0.05 second. From the force vs. displacement curves, work to failure (W; in mJ) was calculated in TestWorks software.

## Wound healing

At 7 weeks of age a 2-mm hole was made in the center of the cartilaginous part of one ear using a metal ear punch. At sacrifice, the whole ear was placed in buffered formalin. To calculate the area of the hole, the ear was first trimmed so


Nature Genetics: doi:10.1038/ng.2644

that it could be placed flat between two slides, and scanned. The hole was delimited and the area calculated using ImageJ.

## Development of a genotyping microarray

### The pilot SNP array

We designed a high density SNP array for the genotyping of the rat HS samples. We focused on the whole genome sampling analysis[10] where genomic complexity is reduced with the selective amplification of restriction fragments. This targeted subset of the genome is used for labeling and hybridization. To this end, we tested several restriction endonucleases. NspI and StyI generated reliable numbers of potential DNA fragments in optimal size.

**Restriction fragments for different endonucleases:**

| Enzyme | Cutting site | Number of Genomic Cutting Sites (millions) | Complexity (>200; <2000 bp) (millions) |
|---|---|---|---|
| NspI | RCATGY | 3.05 | 1,500 |
| StyI* | CC[AT\|TA]GG | 1.25 | 625 |
| HindIII | AAGCTT | 0.8 | 310 |
| XbaI | TCTAGA | 0.7 | 270 |


Nature Genetics: doi:10.1038/ng.2644

From a survey of SNPs from *in-silico* analysis and from re-sequencing data (>6.5 Mio SNPs) we found more than 3 million SNPs that were potentially suited for the array.

**SNP candidates retrieved from sequence information:**

| Resource | Strain | SNP candidates |
|---|---|---|
| Solexa | SHR | 2,364,789 |
| Celera | SD | 2,650,525 |
| STAR | SS, WKY, GK, SHRSP | 128,976 |
| Japan | F344 | 196,812 |
| BCM traces | SR, BB, FHH, SS, LEW, PVG, DA, F344 | 3,634,255 |
| **Total non-redundant** | | 6,644,948 |
| **Chippable** | | 3,209,210 |

**Number of restriction fragments for NspI and StyI:**

| | **NspI** | **StyI** |
|---|---|---|
| Restriction sites | 3,055,638 | 1,245,115 |
| SNP-candidates | 6,644,948 | |
| WGSA-SNP-candidates | 2,118,360 | 929,497 |
| NspI- and StyI*-candidates | **2,497,203** (overlap 550,654) | |
| WGSA-SNP-fragments<br>( > 200 bp ; < 2000 bp) | 987,571<br>= 931 Mbp | 387,902<br>= 413 Mbp |
| Fragments > 10 kb | 4,984 | 25,232 |
| Fragments > 25 kb | 1,817 | 2,592 |

For the design of the rat SNP array, we chose a two phase approach with a pilot array to test candidate SNPs in phase I and a final array with selected high quality SNPs in phase II. Several criteria were assigned for the primary selection of 2.1 million candidate SNPs for the pilot array: SNPs need to reside on a 200-600bp StyI or NspI restriction fragment, the relevant SNP should be within the fragment no closer than 16bp to the cutting sites, no other SNP within 10bp is


Nature Genetics: doi:10.1038/ng.2644

allowed and the 25mer SNP probe set must be unique in the rat genome. Only biallelic SNPs were chosen.

**Example of SNP probes:**


In addition, 7,000 known reference SNPs from our previous SNP genotyping efforts were included[11]. All SNPs were interrogated with 12 probes each.

The selected 2.1 million SNPs were distributed over four pilot arrays, each containing 530K SNPs (including the reference SNPs) which were tested with 95 rat samples. We chose various rat strains for the pilot phase, such as the Brown Norway (BN) reference strain and various inbred and recombinant inbred rat strains, heterozygous rats, 36 technical replicates and 3 mixed samples. The samples were prepared in four batches. Prior to hybridization, the four hybridization mixes of each sample where pooled. In total, 380 arrays were processed.

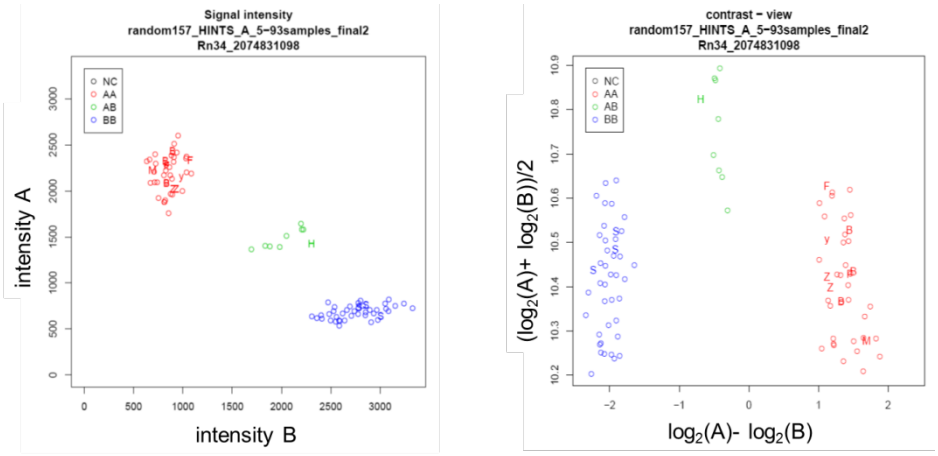
## SNP calling

For the genotype calling of the Pilot Arrays (RAT3 A,B,C and D), we used the programme "apt-probeset-genotype" from the Affymetrix Power Tools (apt-1.10.2). With this programme we used "BRLMM-P" as genotyping method. This method takes normalized data from two alleles across multiple arrays and generates genotypes and confidences for each experiment and each SNP. "BRLMM-P" performs genotyping using three subcomponents: i) generation of tentative genotypes from reference data, ii) creation of cluster models from the tentatively labeled data points and prior information, and iii) the application of the cluster models to yield genotypes.

The first subcomponent generates tentative genotypes for the data points based on a one-dimensional Gaussian mixture model. The likelihood model includes several special features: a Bayesian prior, which allows incorporation of previous information; a penalty for disagreeing with reference data, which allows for errors in training data; and isotonic regression, which prevents clusters from occurring in incorrect orders. These exposed features of the likelihood are "affordances" that simplify training. The algorithm works in "contrast space". The intensity signals for alleles A and B are transformed and plotted in contrast view.


Nature Genetics: doi:10.1038/ng.2644

**Example for signal to contrast space transformation:**




We generated a hints file containing all reference genotype information in order to guide the algorithm. Several parameter-variations for the clustering algorithm were tested to identify optimal settings.

We run a first round of genotyping generating output files for the genotypes, signal intensities, priors, and information about the clustering quality for all 6 probe sets per SNP. We calculated the Fisher Linear Discriminate (FLD) for the cluster model ($FLD_{X\ axis\ values}$ = min((meanAA-meanAB)/sqrt(var),(meanAB-meanBB)/sqrt(var)). FLD describes the discrimination between the clusters of the alleles. We selected all SNPs that showed FLD value greater 4, had less than 3 NoCalls per SNP, minimum1 call for each allele, and the heterozygous cluster within 0.6 data points from the x-axis center. In addition, we used the Het Strength Offset (hetSO) which is the distance between the center of area of the heterozygous and the upper one of the homozygous clusters as quality parameter hetSO ≥ 0.

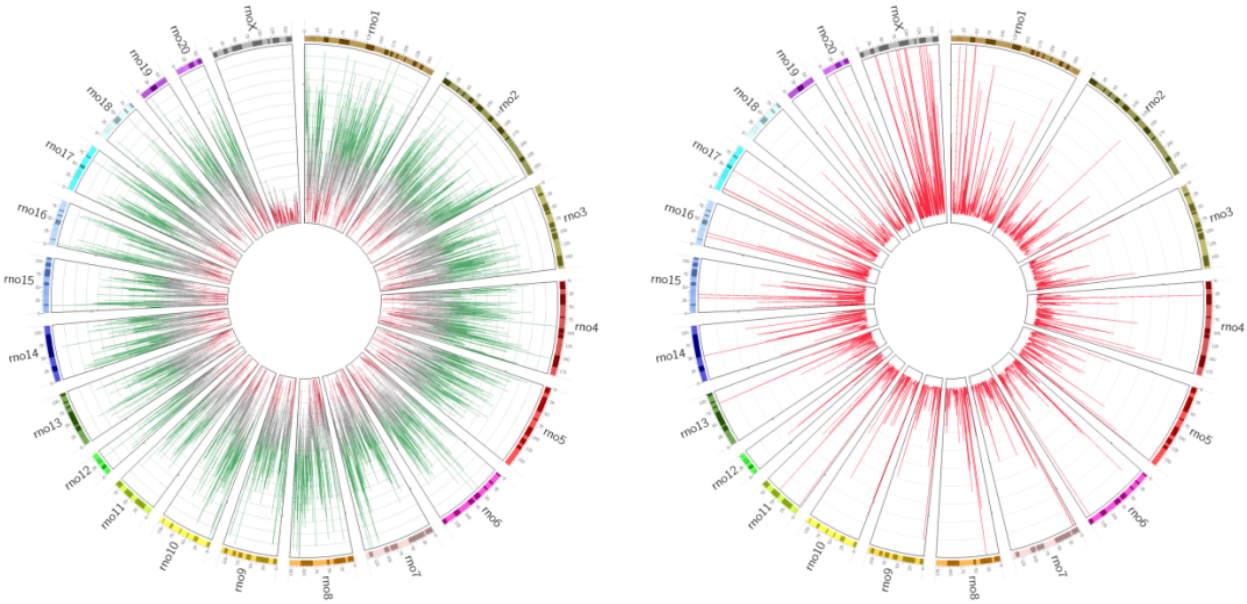
## Final Array Design

With the settings described above, we obtained 803,485 SNPs for which we selected the optimal probe sets for the final array. In total, the RATDIV array consists of almost 6 million probes evenly spaced throughout the rat genome. SNPs are interrogated with one probe set per SNP in 3 or 4 replicates, depending on their FLD. Final cleaning retrieved a ultimate set of 556,358 SNPs with a maximum density of 15 SNPs in a 10kb window. Maximum gap size shown is 1Mb, minimum gap size shown is 50kb. There are 19 larger gaps (1-3.8Mb) on chr1 to chr20 and 12 on chrX, with a maximum of 4.8 Mb.



Nature Genetics: doi:10.1038/ng.2644

**SNP density (left, one scale unit is 1SNP/10kb) and gap size (right, one scale unit is 100kb) of the RATDIV array:**




Additionally, 1 million (1,046,070) copy number (CN) probes have been selected to interrogate known and potential copy number variation (CNV) sites. These probes are located in known CNV in the rat genome[12] and in addition in all constitutively expressed rat exons on basis of ENSEMBL release 56 (ENS56).

## Genotype calling

DNA was extracted in one lab (MDC) while genotyping of the HS samples was conducted in two labs (MDC and CNG). Raw data (cel files) were afterwards collected and analyzed in one batch (MDC). 1407 HS samples were genotyped as well as 198 non-phenotyped parents and 7 founder samples. For genotype calling, we used “apt-probeset-genotype” from the Affymetrix Power Tools apt-1.14.3 with the optimized settings from the RAT3 Pilot arrays.



Nature Genetics: doi:10.1038/ng.2644

## Selection of a set of 500K good quality SNPs available on the RATDIV array

For SNP filtering we applied the following settings: FLD>4, NO of AA calls >9 / AB >3/ BB>9 , >99% call rate per SNP and hetSO ≥ 0. Genotype calls were generated based on the clustering of the contrast values. We noticed that including more samples to the genotyping set improves the overall performance of the SNPs.

We created a list of good working SNPs so that the array can be widely used with this SNP mask and the corresponding annotation. For the analysis, a model file for the RATDIV array can be provided. SNP selection for the RATDIV chip was primarily independent from functional annotation. The final set of high quality SNPs has been annotated using ANNOVAR[13].

**ANNOVAR categories of SNPs:**

| Category | PILOT array | RATDIV array |
|---|---|---|
| Intergenic | 1,518,039 | 396,175 |
| Intronic | 502,486 | 147,806 |
| Exonic | 18,287 | 3,824 |
| Downstream | 14,096 | 3,521 |
| UTR3 | 14,713 | 2,426 |
| Upstream | 10,781 | 2,256 |
| UTR5 | 1,546 | 183 |
| Splicing | 447 | 10 |
| ncRNA_intronic | 285 | 65 |
| Upstream or downstream | 265 | 32 |
| Exonic or splicing | 188 | 50 |
| ncRNA_exonic | 68 | 10 |
| ncRNA_UTR3 | 5 | 0 |
| ncRNA_splicing | 1 | 0 |
| UTR5 or UTR3 | 1 | 0 |
| All | 2,081,208 | 556,358 |


Nature Genetics: doi:10.1038/ng.2644

## Validation of Genotype Calls (on 500K SNPs)

From the HS founder strains (inbred strains BN/SsN, MR/N, BUF/N, M520/N, WN/N, ACI/N, WKY/N, and F344/N) genomic sequence was available which was used to validate the genotype results. For 19,286 SNPs from the 556,358 SNPs we have no genomic sequence data, thus **537,072** SNPs remain for validation. From these SNPs 16,005 show an error in at least one of the 7 HS founder animals and are removed. **521,067** (97%) SNPs remain that pass the filters and are consistent with the sequence data in all strains analysed.

**Number of discrepancies between genotypes derived from the array and genomic sequence (S=sequence, A=array; 0=AA, 1=AB, 2=BB):**

| | ACI_Affy filtered | BUF_Affy filtered | F344_Affy filtered | M520_Affy filtered | MR_Affy filtered | WKY_Affy filtered | WN_Affy filtered |
|---|---|---|---|---|---|---|---|
| S0A1 | 193 | 40 | 117 | 914 | 1,511 | 77 | 592 |
| S0A2 | 88 | 121 | 102 | 113 | 102 | 175 | 85 |
| S1A0 | 53 | 45 | 39 | 40 | 60 | 38 | 43 |
| S1A2 | **1,429** | **1,829** | **1,508** | **1,988** | **1,608** | **2,964** | **1,315** |
| S2A0 | 2 | 2 | 1 | 2 | 2 | 1 | 3 |
| S2A1 | 271 | 52 | 131 | 944 | 956 | 107 | 647 |
| OK | 13,969 | 13,916 | 14,107 | 12,004 | 11,766 | 12,643 | 13,320 |
| ERR | 2,036 | 2,089 | 1,898 | 4,001 | 4,239 | 3,362 | 2,685 |



Nature Genetics: doi:10.1038/ng.2644

**RATDIV Info sheet**

| | SNPs | Probesets | Probes |
|---|---|---|---|
| Affy control SNPs | 1,511 | | |
| RN34 SNPs | 803,484 | 469,853x8P<br>333,631x6P | 5,760,610 |
| WGA SNPs<br>(SNPs on unknown chromosome) | 354 | 313x8P<br>41x6P | 2,750 |
| ENSRNO SNPs<br>(mitochondrial SNPs) | 48 | 48x12P | 576 |
| All SNP Probes | | | **5,763,936** |
| CNV Probes | | | 1,046,069 |
| All not Affy | 803,886 | | |
| All | 805,397 | | 6,810,005 |

After processing over 2,500 samples together and keeping only those SNPs with *FLD* value ≥ 4, call rate per SNP ≥ 99%, minimum calls 9AA/3AB/9BB, and hetSO ≥ 0, we obtained:

| | SNPs that passed QC | SNPs that failed QC | Total number of SNPs |
|---|---|---|---|
| RN34 SNPs | **556,358** | 247,126 | 803,484 |
| WGA SNPs | 64 | 290 | 354 |
| ENSRNO SNPs | 1 | 47 | 48 |
| All | 556,423 | 247,463 | 803,886 |


Nature Genetics: doi:10.1038/ng.2644

Optimized settings for BRLMM-P:

apt-probeset-genotype

--cdf-file RATDIVm520813.CDF

--read-models-brlmmp models.txt

--cel-files celfilelist.txt

--a quant-norm.sketch=50000,pm-only,brlmm-p.CM=1.bins=100.mix=1.bic=2.HARD=3.SB=0.75.KX=0.2.KH=0.3.KXX=-0.1.KAH=-0.1.KHB=-0.1.transform=MVA.AAM=2.0.BBM=-2.0.AAV=0.06.BBV=0.06.ABV=0.06.copyqc=0.00000.wobble=0.05.CSepThr=4.CSepPen=0.1.KYAH=-0.05.KYHB=-0.05.KYAB=-0.1.AAY=10.5.ABY=11.BBY=10.5.copytype=-1.clustertype=2.ocean=0.00001.MS=0.1.hints=1.CP=16.Hok=1

--summaries

--select-probes

--write-models

--no-gender-force

--out-dir RESULTS

## Selection of SNPs for this study

We selected a subset of high quality SNPs to reconstruct the mosaics of progenitor haplotypes in the HS animals. Of the 803,886 SNPs on the RATDIV array, we kept only those whose genotypes for the eight founders were concordant with the genotypes obtained from sequence. We thereby discarded 175K markers. We kept only those SNPs which had three well separated clusters in the (size, contrast) space used by BRLMM-P for genotype calling, which required at least nine homozygote calls of each type and three heterozygote calls, a Fisher Linear Discriminant score greater than 6, and Affymetrix's Heterozygotes Offset value greater than -0.5. We required a call rate higher than 99% (95% on the X chromosome). Finally, based on the genotypes of 96 HS nuclear families, we removed any SNPs with over 4 Mendelian errors (thereby not penalizing a genotyping error in a parental genotype, which would produce apparent Mendelian errors in offspring). 313K markers passed all the previous filters. Finally, we discarded SNPs for which any of the seven genotyped founders was heterozygote, as well as any SNPs that were monomorphic across the seven founders. This yielded 265,551 markers for mapping.



Nature Genetics: doi:10.1038/ng.2644

# Comparison between mixed models and resample model averaging

## Scope of the methods

Mixed models control for false associations by assuming the phenotypic correlation between individuals is proportional to their genome-wide genetic similarity. They control for confounding due to the many loci of small effect that contribute to phenotypic variation. In sesample model averaging techniques such as BAGPHENOTYPE[14], markers compete to explain phenotypic variation in multilocus models. These methods control for false associations arising from long-range correlation with large effect loci.

## Simulation study to compare the performance of the methods

We compared both methods by simulating traits arising both from a large number of loci of small effect (polygenic component) and seven loci with larger effects. Simulations not reported here suggested that the variance components might be inaccurately estimated in mixed models of non-normally distributed phenotypes (binary distribution or negative binomial distribution). Therefore, we compared the methods both on normally and non-normally distributed traits.

We simulated 1,000 normally distributed traits with three components: a polygenic genetic term drawn for a multivariate normal distribution with covariance matrix K and accounting for 20% of phenotypic variation, 7 QTLs simulated from the HAPPY probability matrices, each explaining 5% of phenotypic variation, and uncorrelated errors explaining the remaining 45% of phenotypic variation. Because the polygenic term and the QTLs are correlated, the final effect sizes of the different components can differ substantially from their targets.

We simulated non-normally distributed traits by transforming simulated normally distributed traits, to match the distribution of each of the 19 non-normally distributed trait in our study, by quantile normalisation. We adjusted the variance components used to simulate the phenotypes so that the variance components estimated from the transformed simulated phenotypes matched those observed in each non-normally distributed phenotype. We also matched covariate effects and major QTLs (considered as covariates). We simulated 50 phenotypes for each non-normally distributed measure collected in the HS.

## Comparison of the performance of the methods

Unlike mixed models, resample model averaging does not report QTL associations as p-values, but instead uses inclusion probabilities (the fraction of resampled models containing a given QTL) to quantify the support for a locus. Therefore to make meaningful comparisons between methods we calibrated QTL


Nature Genetics: doi:10.1038/ng.2644

acceptance thresholds so that the QTL false discovery rate (FDR) was the same in each method.

The same sets of simulated phenotypes were analyzed with both methods and QTLs were called. A QTL is here defined as the 6Mb window centered on the interval with the highest negative logP value/inclusion probability, unless this window is truncated because of the presence of an adjacent, more robust QTL. A true association is defined as a QTL encompassing a simulated QTL. For each simulation the detected QTLs were ranked in decreasing order of p-value or inclusion probability, and the number of false associations that ranked above the $k^{th}$ strongest true association was recorded, with $k$ taking values in [1,…,7]. We then pooled simulations to estimate the number of false positive associations found when j true associations have been detected. Since some of the simulated QTLs had a null inclusion probability with resample model averaging, we defined detectable SNPs as those with a non null inclusion probability and compare the FDR of both methods at different proportions of detectable SNPs detected.

# QTL calling

## Inclusion probability thresholds for the resample model averaging analysis

For each non-normally distributed measure and in order to find which inclusion probability threshold corresponded to a 10% FDR, QTLs for the simulated phenotypes were called using inclusion probability thresholds successively between 0.05 and 1 (in steps of 0.05), and the FDR calculated for each threshold. The inclusion probability threshold that led to the FDR closest to 10% was used to call QTLs.

## Significance thresholds for the mixed model analysis

The distribution of the maximum negative log p-value of association under the null hypothesis that no association exist was obtained for each measure by simulating a large number of phenotypes arising from covariates effects, correlated genetic random effects and uncorrelated random errors. The effect sizes of the random terms for these simulations matched those estimated for each phenotype. The genome-wide maximum negative log p-value of association for each simulation was recorded, and an extreme value distribution fitted to these maxima[15].

The logP threshold necessary to achieve a FDR of 10% across all the normally distributed traits was estimated by applying the following procedure for every significance threshold between the $5^{th}$ and $95^{th}$ percentile of the extreme value distributions: call QTLs for each normally distributed phenotype, calculate FDR as the ratio of the number of false positive associations by the number of detected QTLs, where the number of false positive associations is determined by


Nature Genetics: doi:10.1038/ng.2644

the significance threshold (e.g. if the significance threshold is the 65$^{th}$ percentile of the extreme value distribution, there will be a false positive association in 35 out of 100 null genome scans, or 0.35 false associations per scan).

We found that using the 65$^{th}$ percentile of the extreme value distribution for each phenotype ensured a FDR of 10% across all phenotypes.

## Significance thresholds for the merge analysis

200 of the 1,000 simulations used to get significance thresholds for the haplotype-based method were used to get significance thresholds for the SNP-based method and merge analysis. The 65$^{th}$ percentile of the extreme value distribution was used, as for the haplotype-based analysis. Significance thresholds for the SNP-based and merge analyses were thereby obtained for 12 and 20 measures respectively, and we extrapolated significance thresholds for all the other measures from a regression of these thresholds on the thresholds for the haplotype-based method. The $R^2$ of the correlation between the haplotype-based thresholds and the merge analysis thresholds was 0.96 ($p<0.05$) and that for the SNP-based thresholds 0.73 ($p<0.05$). QTLs were then called in the same way as for the haplotype-based analysis.

## Confidence intervals

Confidence intervals for the QTLs mapped using mixed models were calculated by simulating a large number of phenotypes each arising from a single QTL in addition to correlated genetic random effects and uncorrelated errors. The QTLs were simulated with a range of effect sizes, and we binned the simulations according to the logP of the detected QTLs, to obtain confidence intervals for different logP bins. To do so, for each simulation within a bin, the distance between the simulated QTL and the highest interval in the 20Mb window around the QTL was recorded. The distribution of distances obtained this way was used to calculate the 90% confidence interval of the QTL. For those phenotypes mapped using resample model averaging, we report QTLs as 4Mb windows centered on intervals with inclusion probabilities greater than that required to achieve a FDR of 10%.

## QTL calling

After mapping each normally distributed phenotype with mixed models, QTLs were called so that each QTL's confidence interval was defined as detailed above, and centered on the interval with the highest logP. When the boundaries of the confidence intervals of two QTLs were within 2Mb of each other, the two QTLs were merged and assigned the highest logP of the two QTLs.


Nature Genetics: doi:10.1038/ng.2644

# Overlap between sets of QTLs

## Overlap between QTLs mapped in the rat HS and QTLs catalogued in RGD

The QTLs catalogued in the rat genome database (RGD, http://rgd.mcw.edu/) were retrieved from the database on August 9th 2012. For each measure mapped in the HS, we manually identified all the RGD QTLs associated with the exact same measure (although the names could be slightly different). For the analysis, we only considered RGD QTLs smaller than 50Mb, as defined by the start and stop coordinates in the Rnor3.4 rat genome assembly. For each measure, we calculated the number of protein coding and miRNA genes overlapping both the QTLs mapped in the HS and the QTLs catalogued in RGD. To calculate the significance of this overlap, we sampled 1,000 sets of genomic intervals at random so that each set had as many intervals as there are RGD QTLs associated with the measure, and each interval had as many genes as the corresponding RGD QTL. We then calculated the number of genes overlapping both an interval in the random set and an HS QTL. We thereby obtained a distribution of the number of overlapping genes under the null hypothesis of no shared genetic basis between RGD and HS QTLs. The P-value of the overlap between HS and RGD QTLs for a given phenotype was obtained by comparing the number of genes overlapping both RGD and HS QTLs to this null distribution.

## Overlap between the QTLs mapped in the mouse and rat HS

We used the phenotypes and genotypes collected in a mouse HS experiment[15] to compare the extent to which regions associated with the same measure in the mouse HS and rat HS are syntenic. 38 common measures were collected in both heterogeneous stocks. For consistency, we called QTLs for the mouse phenotypes using the same methods as used in the rat HS. To do so, we mapped each measure with mixed models, determined the significance threshold required to have a FDR of 10% across the mouse measures (50th percentile of the extreme value distributions), and called QTLs using the corresponding logP thresholds and 4Mb wide confidence intervals. We then lifted over the mouse QTLs to the rat genome using the UCSC liftover tool (10kb minimum size for the query and target, minimum match of 0.1, and multiple output regions allowed), and calculated for each measure the number of protein coding and miRNA genes overlapping both a rat QTL and a lift-over mouse QTL. To estimate a P-value for each overlap, we then generated 1,000 sets of intervals sampled at random on the mouse genome and so that there were as many intervals as mouse QTLs for the measure and each interval had as many mouse genes as the corresponding QTL. We then lifted over these random intervals to the rat genome, and for each random set calculated the number of genes overlapping both a random interval and a rat QTL. We thereby obtained a null distribution from which we computed the p-value of the overlap between rat and lift-over mouse QTLs for a given phenotype.


Nature Genetics: doi:10.1038/ng.2644

## Guidelines to explore the genome scans and integrated sequence data

The genome scans are available at http://mus.well.ox.ac.uk/gscandb/rat/. These guidelines are also available at http://mus.well.ox.ac.uk/gscandb/rat/help

Select:

- Genome build: **Rattus Norvegicus Rn3.4**
- Population: **HS**
- Plottable Scans: Select phenotypeOfInterest**_new2** to see genome scans for "phenotypeOfInterest" obtained using the mixed models. Select phenotypeOfInterest**_RMIP** to see genome scans obtained with resample model averaging. If multiple Plottable Scans are selected, they will be displayed on separate plots.
- Scan Types. If a Plottable Scan phenotypeOfInterest**_new2** is selected, you will be able to select **additive_haplotype_mapping_logP** for scans based on haplotype mapping and an additive model of allelic effect (i.e. $y_i = \sum_c \beta_c x_{ic} + \sum_s P_{Li}(s)T_{Ls} + u_i + \varepsilon_i$ ). Select **full_haplotype_mapping_logP** for scans based on a full model of allelic effect ($y_i = \sum_c \beta_c x_{ic} + \sum_{s,t} P_{Li}(s,t)(T_{Ls} + T_{Lt}) + u_i + \varepsilon_i$). Select **additive_SNP_mapping_logP** for single (genotyped) marker association. Select **additive_mixed_merge_logP** at the same time as **additive_haplotype_mapping_logP** to compare the variant model (merge analysis) to the haplotype model. Only variants with merge logPs greater than 2 will be displayed. If a Plottable Scan phenotypeOfInterest**_RMIP** is selected, select **additive_RMIP** for scans based on haplotype mapping and an additive effect of allelic effect.
- Click **Genome** to display the 21 chromosomes. Alternatively, select a chromosome in Chromosome menu and click **Region** to display that specific chromosome only. Displaying a single chromosome is also achieved by clicking on a chromosome in the genome view. When a single chromosome is displayed, zoom in on the QTL of interest by pressing and holding the mouse button, dragging it over the QTL and releasing it. Gene annotation becomes available when sufficiently zoomed in. Links to the Rat Genome Database and Ensembl are available for each gene together with its description and coordinates.
- The significance of the thresholds displayed is indicated as the percentile of the extreme value distribution for each phenotype (e.g. 0.5 corresponds to the 50$^{th}$ percentile, and indicates that there will be a false association in 50% of the scans).

Click the **Trait loci table** button to download a table of sequence variants for any window smaller than 6Mb. The columns in the table are: chromosome, start position of the variant, stop position of the variant, reference allele, alleles in the


Nature Genetics: doi:10.1038/ng.2644

eight progenitor strains, negative logarithm of the p-value of association by merge analysis, negative logarithm of the p-value of the partial F-test comparing the variant model and the haplotype model, negative logarithm of the p-value of association with the progenitor haplotypes, variant type (SNP, indel, or type of structural variant), alleles if the variant is a SNP, and if the variant is within 5kb of a gene annotated in Ensembl its annotation (Gene ID, Transcript ID, gene description, and three columns for functional annotation of the variant). All the variants whose alleles for the eight progenitor strains are known (possibly imputed) are in the table. In addition, when the merge logP value is not available for a deleterious SNP or indel (because there are missing values in the variants calls of the progenitor strains), the deleterious SNP or indel will be shown at the top of the table but with no indication of association value by merge analysis.


1. Johannesson, M. *et al.* A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock. *Genome Res* **19**, 150-8 (2009).
2. Lopez-Aumatell, R. *et al.* Effects of environmental and physiological covariates on sex differences in unconditioned and conditioned anxiety and fear in a large sample of genetically heterogeneous (N/Nih-HS) rats. *Behavioral and brain functions : BBF* **7**, 48 (2011).
3. Evans, A.L., Brown, W., Kenyon, C.J., Maxted, K.J. & Smith, D.C. Improved system for measuring systolic blood pressure in the conscious rat. *Medical & biological engineering & computing* **32**, 101-2 (1994).
4. Graham, D., Hamilton, C., Beattie, E., Spiers, A. & Dominiczak, A.F. Comparison of the effects of omapatrilat and irbesartan/hydrochlorothiazide on endothelial function and cardiac hypertrophy in the stroke-prone spontaneously hypertensive rat: sex differences. *Journal of hypertension* **22**, 329-37 (2004).
5. Storch, M.K. *et al.* Autoimmunity to myelin oligodendrocyte glycoprotein in rats mimics the spectrum of multiple sclerosis pathology. *Brain pathology* **8**, 681-94 (1998).
6. Wallström, E.O.T. Rat Models of Experimental Autoimmune Encephalomyelitis. in *Sourcebook of Models for biomedical research* (ed. Conn, P.J.) (Humana Press Inc, Otowa, 2007).
7. Amor, S. *et al.* Identification of epitopes of myelin oligodendrocyte glycoprotein for the induction of experimental allergic encephalomyelitis in SJL and Biozzi AB/H mice. *Journal of immunology* **153**, 4349-56 (1994).
8. Encinas, J.A. *et al.* Identification of genetic loci associated with paralysis, inflammation and weight loss in mouse experimental autoimmune encephalomyelitis. *Int Immunol* **13**, 257-64 (2001).
9. Huang, Y., Jan, K.M., Rumschitzki, D. & Weinbaum, S. Structural changes in rat aortic intima due to transmural pressure. *Journal of biomechanical engineering* **120**, 476-83 (1998).
10. Kennedy, G.C. *et al.* Large-scale genotyping of complex DNA. *Nature Biotechnology* **21**, 1233-7 (2003).
11. Saar, K. *et al.* SNP and haplotype mapping for genetic analysis in the rat. *Nat Genet* **40**, 560-6 (2008).




Nature Genetics: doi:10.1038/ng.2644

12. Guryev, V. *et al.* Distribution and functional impact of DNA copy number variation in the rat. *Nat Genet* **40**, 538-45 (2008).
13. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010).
14. Valdar, W., Holmes, C.C., Mott, R. & Flint, J. Mapping in structured populations by resample model averaging. *Genetics* **182**, 1263-77 (2009).
15. Valdar, W. *et al.* Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat Genet* **38**, 879-87 (2006).



Nature Genetics: doi:10.1038/ng.2644