

Methodology article

Open Access

## A genome-wide SNP panel for mapping and association studies in the rat

Isaac J Nijman, Sylvia Kuipers, Mark Verheul, Victor Guryev and Edwin Cuppen\*

Address: Hubrecht Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

Email: Isaac J Nijman - i.nijman@niob.knaw.nl; Sylvia Kuipers - s.kuipers@niob.knaw.nl; Mark Verheul - m.verheul@niob.knaw.nl; Victor Guryev - v.guryev@niob.knaw.nl; Edwin Cuppen\* - e.cuppen@niob.knaw.nl

\* Corresponding author

Published: 25 February 2008

Received: 4 December 2007

BMC Genomics 2008, 9:95 doi:10.1186/1471-2164-9-95

Accepted: 25 February 2008

This article is available from: <http://www.biomedcentral.com/1471-2164/9/95>

© 2008 Nijman et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The laboratory rat (*Rattus norvegicus*) is an important model for human disease, and is extensively used for studying complex traits for example in the physiological and pharmacological fields. To facilitate genetic studies like QTL mapping, genetic markers that can be easily typed, like SNPs, are essential.

**Results:** A genome-wide set of 820 SNP assays was designed for the KASPar genotyping platform, which uses a technique based on allele specific oligo extension and energy transfer-based detection. SNPs were chosen to be equally spread along all chromosomes except Y and to be polymorphic between Brown Norway and SS or Wistar rat strains based on data from the rat HapMap EU project. This panel was tested on 38 rats of 34 different strains and 3 wild rats to determine the level of polymorphism and to generate a phylogenetic network to show their genetic relationships. As a proof of principle we used this panel to map an obesity trait in Zucker rats and confirmed significant linkage (LOD 122) to chromosome 5: 119–129 Mb, where the leptin receptor gene (*Lepr*) is located (chr5: 122 Mb).

**Conclusion:** We provide a fast and cost-effective platform for genome-wide SNP typing, which can be used for first-pass genetic mapping and association studies in a wide variety of rat strains.

### Background

The laboratory rat has proven to be an important model organism for human disease, physiology, immunology and pharmacology [1,2]. Over the years, selective breeding for disorders has resulted in the establishment of more than 500 inbred lines that allows the study of factors involved in human multi-factorial diseases under controlled experimental conditions [3]. Genetic mapping and identification of components that affect disease-related

phenotypes in the rat can be extremely useful to improve insight in the analogous human syndromes [4].

Traditionally, QTL mapping studies in rat (*Rattus norvegicus*) have been limited by the availability of markers and often focused marker development was needed to successfully link genomic regions to a specific trait (for example [5,6]). Genome information has increased dramatically with technological advances [7] and consequently the number of candidate markers has increased significantly.

However, due to the phylogenetic relationship of rat inbred strains, large haplotype blocks can be found grouping several SNPs together. The EU framework 6 programme 'STAR, a SNP and haplotype map for the rat', set out to identify these regions [8]. Combined with new genotyping methods, cheap and versatile genome-wide genotyping has come within reach [9]. Here, we describe the design and application of a panel of 820 rat SNPs for the versatile and fast KASPar system (KBiosciences, UK), which uses a competitive allele-specific PCR combined with a FRET quenching reporter oligo. For each SNP, three unlabeled oligos need to be synthesized, but the assay runs as a standard PCR. After amplification using genomic DNA as a template, the fluorophore signals are determined and genotypes are automatically determined. The SNP panel and genotyping information for more than 30 commonly used rat inbred strains reported here forms a versatile basis for the design of first round genetic mapping and association studies in a wide range of rat strains.

## Results and Discussion

### Selection of SNP candidates for assay setup

Candidate SNPs were selected from various sources (384 and 478 from [8,10] respectively) with a preference for polymorphism between BN and SS or Wistar and an equal physical distribution in the genome. This set was validated in duplo on BN ( $n = 2$ ), Wistar ( $n = 2$ ) and BN-Wistar mixed DNA ( $n = 4$ ) to verify that both alleles could be amplified and heterozygotes are easily discerned from either homozygous genotype. From the originally designed 862 assays, 820 worked reliably (Fig 1) – i.e. showing clear clustering of genotypes- and subsets from these were used for further analyses. The conversion rate of 95% from candidate polymorphism to robust assay is relatively high for a SNP genotyping platform. It should be noted that we did not perform any assay optimization in terms of oligo design or experimental conditions (e.g. numbers of cycles or annealing temperatures) and each assay was run with universal parameters.

### Genotyping

All 820 assays were typed in duplo on a panel of BN ( $n = 2$ ), Wistar ( $n = 2$ ), 31 other widely used inbred rat strains that are commercially available, 5 animals from a wild-derived rat strain and 3 wild rats [10]. After removing failed assays, 34,398 duplo genotyping assays remained. From these assays, 49 genotype scores were discrepant between the duplos (0.14%, clustered in 40 SNPs). Another 20 SNPs indicated some heterozygosity while only homozygotes were expected in inbred strains. Additionally, in 1010 cases we could not reliably score one of the duplos (2.9%, clustered in 472 SNPs). In the final dataset, these scores were removed and the final genotype was dictated by the single assay that produced a clear gen-

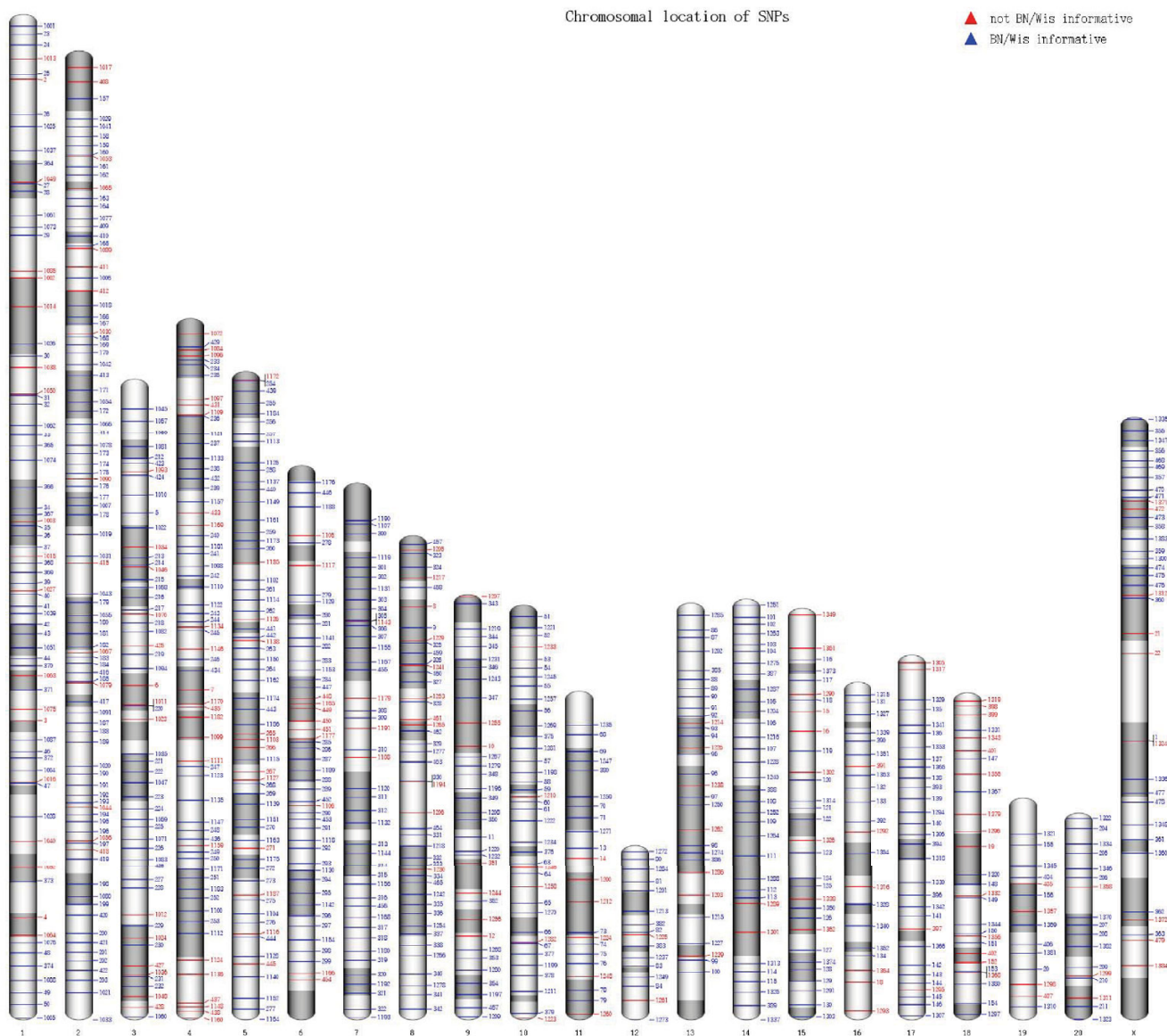
otype. Overall, 347 SNPs produced perfect duplo genotypes for all the strains without any uncertainty.

As expected, the genotyping clearly distinguished BN from SS or Wistar (Fig. 2) since SNP design was biased to be polymorphic between these two strains. While the outbred Wistar strain contains much more variable positions, lack of detectable polymorphisms in a few animals can obscure a potential informative marker. Therefore we selected two genome-wide sets of markers. The first panel consists of a set of 632 markers that are polymorphic between BN and the tested Wistar animals. The other markers are not informative in our BN/Wistar crosses, but were chosen to cover big chromosomal gaps where no informative markers were available for BN/Wistar. All the other rat strains show an intermediate pattern with a median number of 231 polymorphic markers for any pairwise strain combination, indicating that this panel is also informative for mapping studies using a wide range of rat strains. Figure 3 shows a pairwise comparison of allele differences between strains. In this case, heterozygote markers are included since they can be informative. In our follow-up study, only the markers with homozygous alleles were used. All genotyping data, including information on the genomic locus and oligonucleotide sequences, can be explored for any arbitrary pairwise strain combination [11].

The hierarchical clustering tree (Fig. 4) shows the relationships between the strains and concurs with known genetic background [10] and previous studies [8]. The domesticated wild-derived rats from Groningen (wild II,  $n = 5$ ) are much more related to each other than the wild rats caught in Utrecht, Amsterdam and Beers, which is most likely due to genetic drift in the colony which is now maintained in the lab for over 30 years.

### Phylogenetic analysis

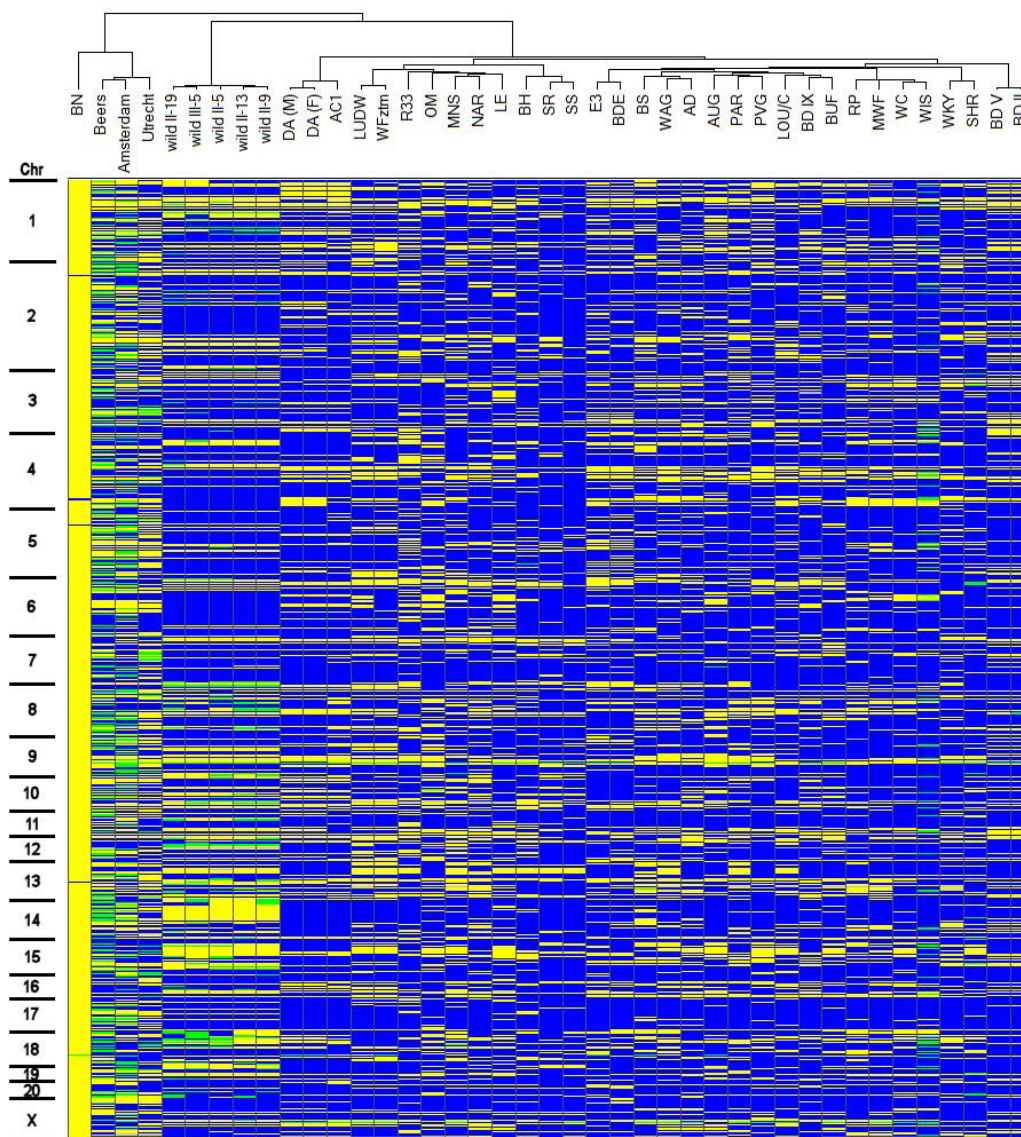
The genotypes of a genome wide SNP panel provide a good data set to infer a phylogenetic structure for this panel of inbred strains. At present, this structure is only available based on a small amount of SNPs [10] or microsatellites [12,13]. The general topology of the hierarchical clustering of Figure 2 is supported by a separate phylogenetic analysis resulting in a Neighbor-Joining tree (Fig. 4). From this tree it is clear that the older nodes are only weakly supported, which is not surprising since a combination of different strains has been used to generate the current inbred strains. Therefore, it is unknown how much each parental strain has contributed to the current strains and a high level of reticulation can be expected. A NeighborNet analysis (Fig. 5) indeed confirms this complex ancestral history and provides a more unbiased view of phylogenetic clusters. From both figures the large branch length towards BN is obvious. This probably



**Figure 1**  
**Chromosomal locations of SNPs on the RSGC3.4 rat genome assembly.** The image was drawn using the idiographica webtool [22]. Giemsa staining is set as chromosomal shading background. Markers indicated in blue were informative between our BN and Wistar animals, and red markers were not, but showed polymorphism in other strains. SNPs in pairwise combinations of strains can be explored at <http://cascad.niob.knaw.nl>.

reflects the ascertainment bias because SNPs are selected from a comparison with the published genome sequence of the Brown Norway rat [7]. However, similar studies using microsatellite markers, which are less sensitive to this bias, also showed that BN is relatively far diverged from the other rat inbred lines [12,13]. In general, a number of clusters can be distinguished in the unrooted Neighbor-net: 1: SS-like, 2: LUDW-like, 3: Wistar-like, 4: BD-like, 5: WKY/SHR, 6: DA-like and the group contain-

ing BN and the wild rats. Most groups are formed by strains that are known to be closely related. For instance in the BD cluster, BDE is an inbred strain crossed from the BD group and E3. Additionally, WKY and SHR are closely related [3] and indeed cluster tightly together. The strains clustering in the Wistar group have known Wistar background [3].



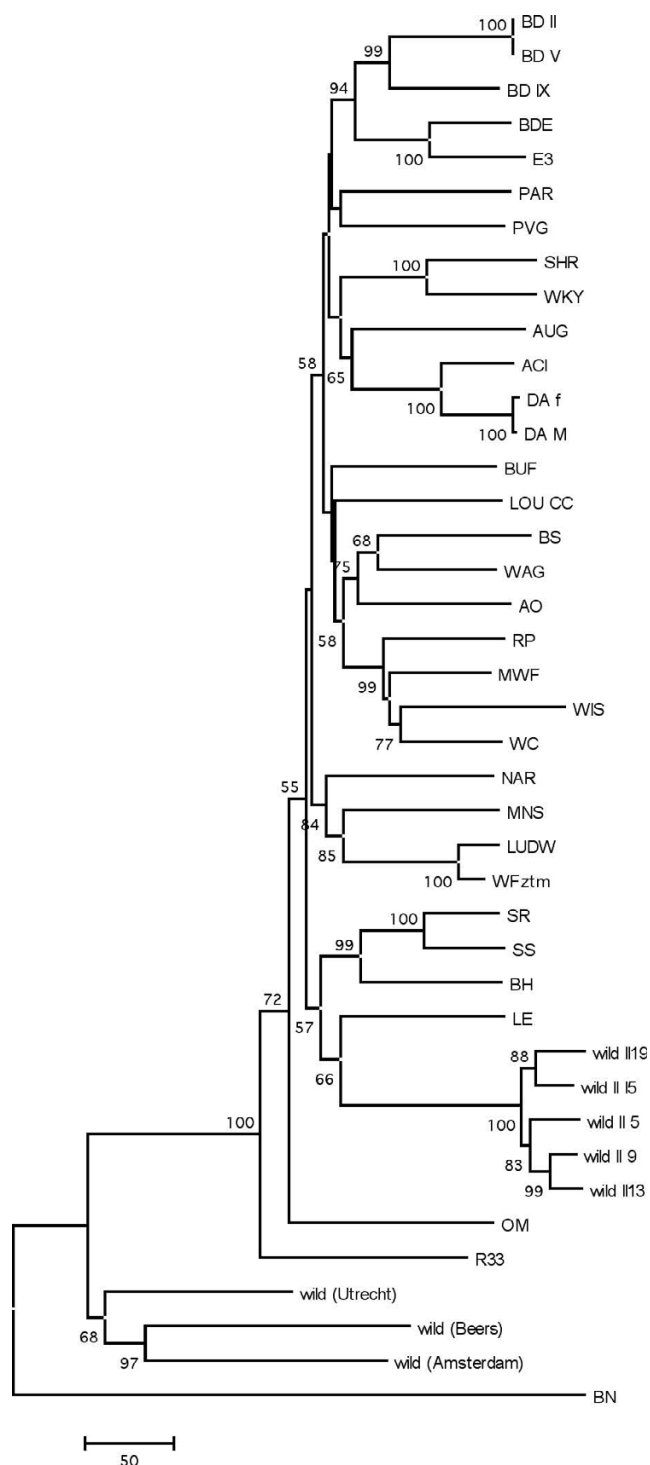
**Figure 2**  
**Heat map of genotyping results of 820 markers in 33 rat strains and 3 wild rats.** Markers are sorted by chromosome and the position of the maker on that chromosome. Yellow indicates BN alleles, blue indicates polymorphic alleles as compared to BN and green positions indicate heterozygous genotypes. A hierarchical cluster tree, based on Wards algorithm is shown on top.

**Mapping of the obese trait in Zucker rats**

As a proof of principle for a mapping study, we used our validated SNP panel in a cross with the Zucker rat strain (originally a cross between Sherman and Merck 13 M strains [14]), a strain which is commonly used as a model in obesity research [15]. For this study, animals from the Zucker strain were outcrossed with Brown Norway and F1 animals were backcrossed with Zucker rats. The resulting BC animals were selected for genotyping based on their obese appearance at 3–4 weeks of age. As our SNP panel

was developed to contain polymorphic markers between BN and Wistar or SS (which shares many alleles with Wistar rats [10]), a large number of informative markers was expected. Indeed, 205 out of 547 markers tested were found to be polymorphic and these were genotyped on the 239 animals from the Zucker – BN backcross. After genotyping, the Bn/Wistar mix controls showed a complete heterozygous profile indicating that all alleles were correctly amplified and detected. Genotype calling was successful for 99% of the assays (48,489 out of 48,995).





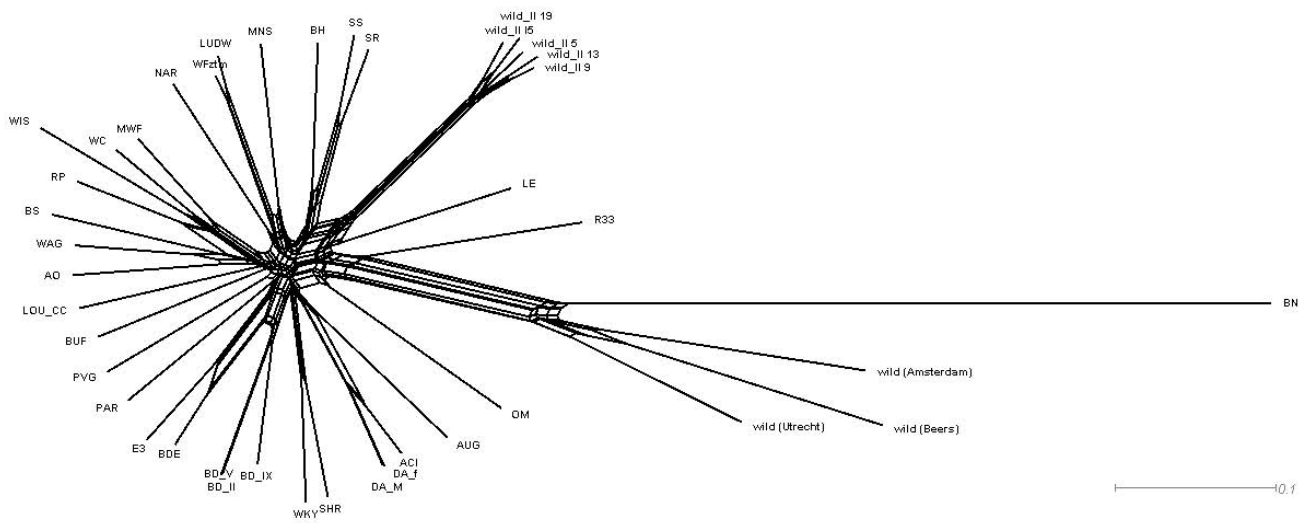
**Figure 4**  
**A Neighbor-joining phylogenetic tree of rat strains based on uncorrected p-distances of the genotypes of complete set of 820 markers.**

Ztm (67981), BH/Ztm (67989), BN/Crl (60985), BS/Ztm (68008), BUF/SimRijHsd (60986), DA/Ztm (60997), E3/Ztm (61013), LE/Ztm (60991), LOU/CZtm (68079), LUDW/OlaHsd (68082), MNS/Ztm (60992), MWF/Ztm (68099), NAR/Ztm (737969), OM/Ztm (70452), PAR/Ztm (737940), PVG/OlaHsd (61006), R33/Ztm, RP/Aeur-RijHsd (68127), SHR (61000), SR/JrHsd (70453), SS/JrHsd (69369), WAG/RijHsd (61008), WC/Ztm, WF/Ztm (61007), WKY/Ztm (61103), outbred Crl:Wistar (10044) and wild/Gro (5 animals wild\_11 5–15). Additionally, 3 wild rats caught in Utrecht (wild1/Hubr; 1625284), Amsterdam, and Beers in the Netherlands were analysed.

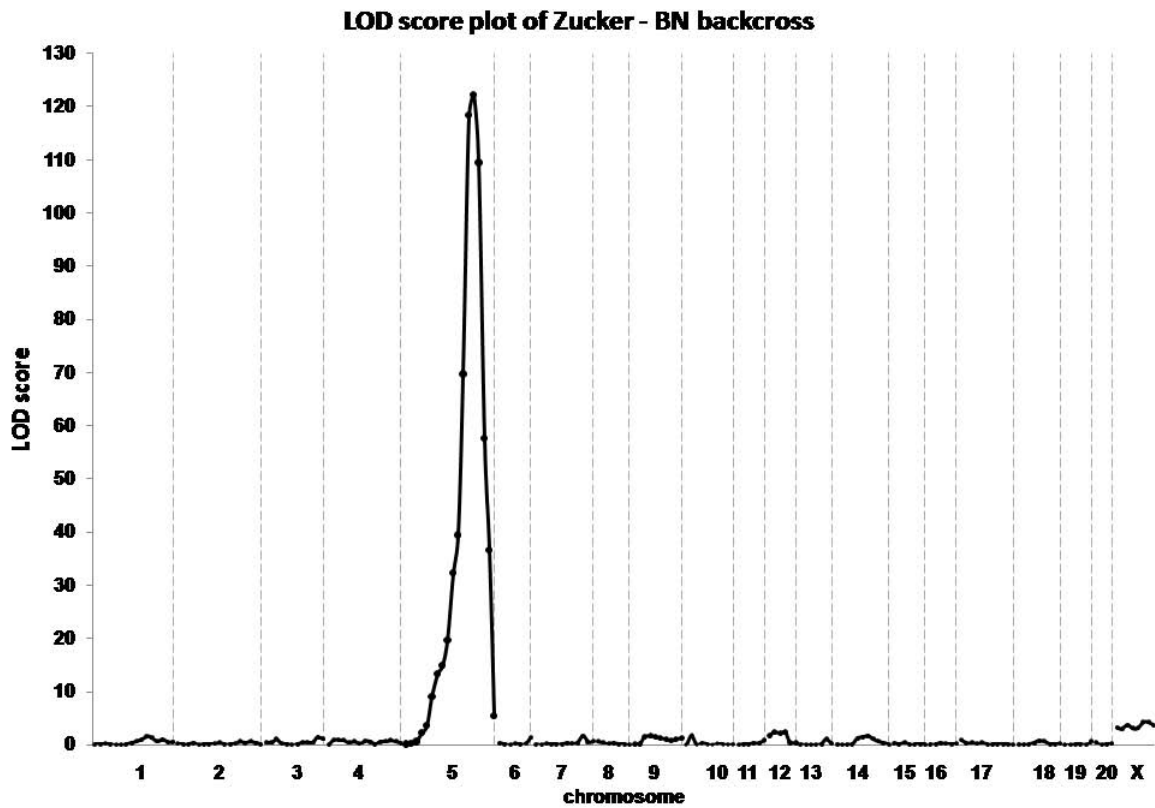
For the trait mapping analysis we obtained DNA from 239 BC rats which were selected for an obesity phenotype by visual inspection at an age of 3–4 weeks. These animals were derived from a backcross of Zucker/BN F1 heterozygotes with Zucker animals.

**SNP selection and assay design**

A Perl script was written to select SNPs from genome and rat HapMap data, evenly spread over the chromosomes and polymorphic between BN and Wistar or SS (Fig 1. [11] and Additional file 1). For a set of 862 SNPs, primers were designed for KASPar genotyping using a tool provided by KBiosciences [18] based on the SNP locus sequence (about 50 nt flanking each side of the SNP are required for the design). The output provides sequence information for two allele-specific oligonucleotides of about 40 nt in length and 1 common oligonucleotide of about 20 nt in length, all of which are standard unmodified and unlabelled oligonucleotides. Detailed information on every marker can be found in Additional file 1 and [11]. The three oligonucleotides for each assay were dissolved in 10 mM Tris-HCl (pH 8) to a 100 µM concentration, mixed together as a SNP assay mix (12 µl AS1 + 12 µl AS2 + 30 µl CP + 46 µl; Tris-HCl pH 8) and 2 µl aliquots were distributed into individual wells of 384 well plates by a Tecan Robot (Genesis RSP200 liquid handling workstation including an integrated 96-channel pipetting head TMO96). Assay plates were frozen at -20°C until use. Each SNP was typed in a total volume of 4 µl in the following reaction mixture: 6 ng DNA, 22 mM MgCl<sub>2</sub>, KTAq, 1 µl 4× reaction mix, 2 µl pre-plated assay mix according to the manufacturer's guidelines (KBiosciences). Amplification was performed in Applied Biosystems GeneAmp 9700 thermocyclers running the following program: 94°C - 15' then 20 cycles of 94°C-10", 57°C-5" and 72°C-10", followed by 18 cycles of 94°C-10", 57°C-20" and 72°C-40". Fluorescence scanning of the reactions was done in a BMG labtech Pherastar scanner and the results were interpreted by the KlusterCaller 1.1 software (KBiosciences). Per 384 well plate, all SNPs are amplified for a single individual and afterwards all data for each locus is regrouped



**Figure 5**  
**A NeighborNet phylogenetic network of rat strains based on uncorrected p-distances of the genotypes of complete set of 820 markers.**



**Figure 6**  
**Linkage analysis graph based on the genotyping results in the Zucker/BN backcross population using 205 informative makers. A highly significant LOD score for the leptin receptor locus on chromosome 5 is obtained.**

for all samples by a custom Perl script before interpretation by KlusterCaller.

### Phylogenetic analysis

The phylogenetic analysis of genotypes of the rat strains was performed using the MEGA package [19] and Splitstree [20]. The Neighbour-joining tree and NeighborNet diagram were calculated based on uncorrected p distances. A heat map of SNP data and the hierarchical clustering (Ward's method) tree were drawn by Spotfire Decisionsite 8.2.1 (Tibco, US).

### Mapping the obesity trait

A custom Perl script was written to calculate LOD scores for the obesity trait in the Zucker population. The method used was based on maximum likelihood estimates of the recombination fraction [21]. This script is available upon request.

### Authors' contributions

IN designed the SNP panel, carried out analyses and wrote the manuscript. MV and SK performed Kaspar typings. VG designed part of the panel and assisted in bioinformatic analyses. EC initiated and supervised the project and contributed to the manuscript.

All authors read and approved the final manuscript.

### Additional material

#### Additional file 1

Supplementary table. All validated KASPar rat SNP markers designed on the BN genome sequence (RGSC 3.4). The map id refers to the graphical representation in Figure 1 (main text).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-95-S1.xls>]

### Acknowledgements

The mapping of the Zucker obesity trait was done in collaboration with S. Hansen, C. Warden and J. Stern, University of California, Nutrition department, Davis CA, USA. We greatly appreciate the assistance of KBiosciences, UK for advice and help setting up the KASPar genotyping platform and oligo design.

### References

- Gill TJ 3rd, Smith GJ, Wissler RW, Kunz HW: **The rat as an experimental animal.** *Science* 1989, **245(4915)**:269-276.
- Jacob HJ, Kwitek AE: **Rat genetics: attaching physiology and pharmacology to the genome.** *Nat Rev Genet* 2002, **3(1)**:33-42.
- Rat Genome Database** [<http://rgd.mcw.edu/strains/>]
- Lazar J, Moreno C, Jacob HJ, Kwitek AE: **Impact of genomics on research in the rat.** *Genome Res* 2005, **15(12)**:1717-1728.
- Dukhanina OI, Dene H, Deng AY, Choi CR, Hoebee B, Rapp JP: **Linkage map and congenic strains to localize blood pressure QTL on rat chromosome 10.** *Mamm Genome* 1997, **8(4)**:229-235.
- Otsen M, den Bieman M, Kuiper MT, Pravenec M, Kren V, Kurtz TW, Jacob HJ, Lankhorst A, van Zutphen BF: **Use of AFLP markers for gene mapping and QTL detection in the rat.** *Genomics* 1996, **37(3)**:289-294.
- Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, Scherer S, Scott G, Steffen D, Worley KC, Burch PE, Okwuonu G, Hines S, Lewis L, DeRamo C, Delgado O, Dugan-Rocha S, Miner G, Morgan M, Hawes A, Gill R, Celera, Holt RA, Adams MD, Amanatides PG, Baden-Tillson H, Barnstead M, Chin S, Evans CA, Ferreira S, Foster C, Glodek A, Gu Z, Jennings D, Kraft CL, Nguyen T, Pfannkoch CM, Sitter C, Sutton GG, Venter JC, Woodage T, Smith D, Lee HM, Gustafson E, Cahill P, Kana A, Doucette-Stamm L, Weinstock K, Fechtel K, Weiss RB, Dunn DM, Green ED, Blakesley RW, Bouffard GG, De Jong PJ, Osoegawa K, Zhu B, Marra M, Schein J, Bosdet I, Fjell C, Jones S, Krzywinski M, Mathewson C, Siddiqui A, Wye N, McPherson J, Zhao S, Fraser CM, Shetty J, Shatsman S, Geer K, Chen Y, Abramson S, Nierman WC, Havlak PH, Chen R, Durbin KJ, Egan A, Ren Y, Song XZ, Li B, Liu Y, Qin X, Cawley S, Worley KC, Cooney AJ, D'Souza LM, Martin K, Wu JQ, Gonzalez-Garay ML, Jackson AR, Kalafas KJ, McLeod MP, Milosavljevic A, Virk D, Volkov A, Wheeler DA, Zhang Z, Bailey JA, Eichler EE, Tuzun E, Birney E, Mongin E, Ureta-Vidal A, Woodwark C, Zdobnov E, Bork P, Suyama M, Torrents D, Alexandersson M, Trask BJ, Young JM, Huang H, Wang H, Xing H, Daniels S, Gietzen D, Schmidt J, Stevens K, Vitt U, Wingrove J, Camara F, Mar Alba M, Abril JF, Guigo R, Smit A, Dubchak I, Rubin EM, Couronne O, Poliakov A, Hubner N, Ganten D, Goesele C, Hummel O, Kreitler T, Lee YA, Monti J, Schulz H, Zimdahl H, Himmelbauer H, Lehrach H, Jacob HJ, Bromberg S, Gullings-Handley J, Jensen-Seaman MI, Kwitek AE, Lazar J, Pasko D, Tonellato PJ, Twigger S, Ponting CP, Duarte JM, Rice S, Goodstadt L, Beatson SA, Emes RD, Winter EE, Webber C, Brandt P, Nyakatura G, Adetobi M, Chiaromonte F, Elnitski L, Eswara P, Hardison RC, Hou M, Kolbe D, Makova K, Miller W, Nekrutenko A, Riemer C, Schwartz S, Taylor J, Yang S, Zhang Y, Lindpaintner K, Andrews TD, Caccamo M, Clamp M, Clarke L, Curwen V, Durbin R, Eyras E, Searle SM, Cooper GM, Batzoglu S, Brudno M, Sidow A, Stone EA, Venter JC, Payseur BA, Bourque G, Lopez-Otin C, Puente XS, Chakrabarti K, Chatterji S, Dewey C, Pachter L, Bray N, Yap VB, Caspi A, Tesler G, Pevzner PA, Haussler D, Roskin KM, Baertsch R, Clawson H, Furey TS, Hinrichs AS, Karolchik D, Kent WJ, Rosenbloom KR, Trumbower H, Weirauch M, Cooper DN, Stenson PD, Ma B, Brent M, Arumugam M, Shetye N, Copley RR, Taylor MS, Riethman H, Mudunuri U, Peterson J, Guyer M, Felsenfeld A, Old S, Mockrin S, Collins F: **Genome sequence of the Brown Norway rat yields insights into mammalian evolution.** *Nature* 2004, **428(6982)**:493-521.
- STAR: A SNP and Haplotype map for the rat.** [<http://www.snp-star.eu/>]
- Meaburn E, Butcher LM, Schalkwyk LC, Plomin R: **Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans.** *Nucleic Acids Res* 2006, **34(4)**:e27.
- Smits BM, Guryev V, Zeegers D, Wedekind D, Hedrich HJ, Cuppen E: **Efficient single nucleotide polymorphism discovery in laboratory rat strains using wild rat-derived SNP candidates.** *BMC Genomics* 2005, **6**:170.
- CASCAD SNPview** [<http://cascad.niob.knaw.nl/snpview/>]
- Mashimo T, Voigt B, Tsurumi T, Naoi K, Nakanishi S, Yamasaki K, Kuramoto T, Serikawa T: **A set of highly informative rat simple sequence length polymorphism (SSLP) markers and genetically defined rat strains.** *BMC Genet* 2006, **7**:19.
- Thomas MA, Chen CF, Jensen-Seaman MI, Tonellato PJ, Twigger SN: **Phylogenetics of rat inbred strains.** *Mamm Genome* 2003, **14(1)**:61-64.
- Zucker LM, Zucker TF: **Fatty, a new mutation in the rat.** *J Hered* 1961, **52**:275-278.
- Stevenson FT, Wheeldon CM, Gades MD, van Goor H, Stern JS: **Hyperphagia as a mediator of renal disease initiation in obese Zucker rats.** *Obes Res* 2001, **9(8)**:492-499.
- Phillips MS, Liu Q, Hammond HA, Dugan V, Hey PJ, Caskey CJ, Hess JF: **Leptin receptor missense mutation in the fatty Zucker rat.** *Nat Genet* 1996, **13(1)**:18-19.
- Truett GE, Bahary N, Friedman JM, Leibel RL: **Rat obesity gene fatty (fa) maps to chromosome 5: evidence for homology with the mouse gene diabetes (db).** *Proc Natl Acad Sci U S A* 1991, **88(17)**:7806-7809.
- Primer-Picker** [<http://kbioscience.co.uk/primer-picker/>]



19. Kumar S, Tamura K, Nei M: **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.** *Brief Bioinform* 2004, **5(2)**:150-163.
20. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies.** *Mol Biol Evol* 2006, **23(2)**:254-267.
21. Morton NE: **Sequential tests for the detection of linkage.** *Am J Hum Genet* 1955, **7(3)**:277-318.
22. Kin T, Ono Y: **Idiographica: a general-purpose web application to build idiograms on-demand for human, mouse and rat.** *Bioinformatics* 2007, **23(21)**:2945-2946. Epub 2007 Sep 24.

Publish with **BioMed Central** and every scientist can read your work free of charge

*"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."*

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

