A genetic and genomic analysis identifies a cluster of genes associated with hematopoietic cell turnover.

Short title: A locus affecting hematopoietic cell proliferation

Gerald de Haan*, Leonid V. Bystrykh*, Ellen Weersing*, Bert Dontje*, Hartmut Geiger, Natalia Ivanova§, Ihor R. Lemischka§, Edo Vellenga¶, Gary Van Zant*\n
*Department of Stem Cell Biology, University of Groningen, A. Deusinglaan 1, 9713 AV Groningen, the Netherlands, Division of Hematology/Oncology, Blood and Marrow Transplant Program, Markey Cancer Center, University of Kentucky, 800 Rose St., Lexington, KY, §Department of Molecular Biology, Princeton University, Princeton, NJ, ¶Department of Hematology, University Hospital Groningen, Groningen, the Netherlands, 55Department of Physiology, University of Kentucky Medical Center, 800 Rose St., Lexington, KY.

*To whom correspondence should be addressed. E-mail: g.de.haan@med.rug.nl

This work was supported by the Dutch Cancer Society (RUG 2000-2183), a fellowship from the Royal Netherlands Academy of Arts and Sciences (KNAW) to GdH, NIH grant AG16653 to GVZ, and a fellowship of the Deutsche Akademie der Naturforscher Leopoldina funded by the Bundesministerium fuer Bildung und Forschung to HG.

Number of words in abstract: 222  Number of characters in paper: 4264
ABSTRACT

Hematopoietic stem cells from different strains of mice vary widely with respect to their cell cycle activity. In the present study we used complementary genetic and genomic approaches to identify molecular pathways affecting this complex trait. We identified a major quantitative trait locus (QTL) associated with variation in cell proliferation in C57BL/6 and DBA/2 mice to a 10 cM region on chromosome 11. A congenic mouse model confirmed that a genomic interval on chromosome 11 in isolation confers the proliferation phenotype. To detect candidate genes we performed subtractive hybridizations and gene arrays using cDNA from highly enriched stem cells from parental strains. Intriguingly, a disproportionate number of differentially expressed genes mapped to chromosome 11 and, more specifically, these transcripts occurred in three distinct clusters. The largest cluster co-localized exactly with the cell cycling QTL. Such clustering suggested the involvement of genetic variation that affects higher order chromosomal organization. This hypothesis was reinforced by the fact that differentially expressed genes mapped to recombination “coldspots”, as a consequence of which clustered genes are collectively inherited. These findings suggest the functional interdependence of these closely linked genes. Our data are consistent with the hypothesis that this isolated cell cycle QTL does not result from a mutation in a single gene, but rather is a consequence of variable expression of a collection of highly linked genes.
INTRODUCTION

Lifelong replenishment of short-lived circulating blood cells is ensured by a relatively small population of lympho-hematopoietic stem- and progenitor cells residing in the bone marrow. Mechanisms regulating proliferation of these primitive cells during steady state blood cell production are poorly understood. It has long been assumed that most hematopoietic stem cells are quiescent for prolonged periods of time and do not actively contribute to mature blood cell production. However, this concept has been challenged by in vivo bromodeoxyuridine labeling studies that indicated that in fact most, if not all, stem cells in the mouse will divide at least once during a period of 3-4 weeks, suggesting that all stem cells participate in steady state blood cell production. In addition to a plethora of cell-extrinsically acting cytokines, such as the ligands for c-kit, flt3/flk2, and c-mpl, that have been demonstrated to affect kinetics of stem- and progenitor cells, superimposed on such a growth factor network intrinsic genetic factors play an important role in this process as well. We have shown by analyzing various inbred strains of mice that the rate of progenitor cell cycling varies dramatically from strain to strain. Intriguingly, we found that strains with rapidly cycling stem- and progenitor cells (e.g. AKR/J, C3H, DBA) invariably had a shorter mean lifespan than strains with more slowly proliferating cells (B6, 129/J). In order to assess whether these two traits were causally related, we have mapped quantitative trait loci (QTL) that were associated with variation in cell cycling in a panel of BXD recombinant inbred mice and compared these with loci associated with mean lifespan. Indeed two loci, occurring on chromosome 7 and 11 respectively, showed suggestive linkage with both traits.
The heritable component of quantitative traits, such as variation in cell cycling, is thought to result from the additive and epistatic interactions of multiple genetic loci. The ensuing non-Mendelian segregation pattern has significantly hampered the identification of individual genes associated with quantitative traits. Strategies such as the production of congenic or consomic strains, carrying only the QTL or chromosome of interest, have been proposed to circumvent this problem. In the present study we followed this approach, and by employing a variety of complementary genetic and genomic approaches we have searched for molecular pathways affecting hematopoietic progenitor cell proliferation.
MATERIALS AND METHODS

Mice: 8 to 12 weeks old female recombinant inbred BXD-33, 34, 35, 38, 39, 40 and 42/TyJ were purchased from The Jackson Laboratory (Bar Harbor, ME). C57BL/6 (B6), DBA/2 (DBA) and B6D2F1 (BDF1) mice were all from Harlan (Harlan Nederland BV, Horst, the Netherlands).

Generation of backcross animals and the use of a marker-assisted selection protocol to produce congenic mice. To produce N2 backcross animals, male BDF1 mice were crossed with female B6 mice at the Central Animal Facility of the University of Groningen. To generate congenic mice we used a marker-assisted selection protocol 18, and selected male carriers of DBA alleles at the region of interest on chromosome 11 that otherwise showed the lowest frequency of DBA alleles throughout the genome, using 57 MapPairs primers (Research Genetics, Huntsville, AL). For genetic linkage analyses mice were genotyped with an aditional 15 markers on chromosome 11 (D11Mit226, -135, -20, -154, -131, -349, -278, -177, -212, -285, -67, -132, -224, -181, and -338). PCR conditions were according to the manufacturer’s suggestions, and DNA was resolved on 3.5-4% agarose gels (GIBCO-BRL).

All mice were kept in sterile microisolators and were fed sterile food and water ad libitum.

Progenitor and stem cell assays: The percentage of hematopoietic progenitors in S-phase was determined using standard semisolid colony forming units-granulocyte/macrophage (CFU-GM) assays, supplemented with recombinant rat Stem Cell Factor (a gift from Amgen Inc., Thousand Oaks, CA) and recombinant murine GM-CSF (a gift from Behringwerke AG, Marburg, Germany). Prior to their use in the
progenitor assay bone marrow cells were incubated with or without the S-phase specific cytotoxic drug hydroxyurea (200 µg/ml, Sigma) for 1 hr. Cells were washed twice and CFU-GM assays were subsequently performed using standard procedures. Cultures were evaluated after 7 days. The fraction of progenitors that was killed by hydroxyurea incubation was calculated and was considered to have been in the S-phase of the cell cycle. Cobblestone area forming cell (CAFC) assays to quantify progenitor and stem cell frequencies, were performed exactly as described previously 13. Both assays were chosen for consistency and to allow comparison with our previously published data 12, 13, 15.

**Competitive repopulation assay and chimerism analysis:** Recipient C57BL/6.SJL.Ptprca (Ly5.1+/CD45.1+) were conditioned with a sublethal dose of 4 Gy total body γ-irradiation administered in one single dose 24 hrs. prior to transplantation. Recipient mice received a dose of 3 x 10^6 cells unfractionated bone marrow cells, derived from wildtype B6 (n=6) or congenic B6.DBA^scp2 (n=6) donors (both strains are Ly5.2+/CD45.2+). To determine the percentage donor blood cell contribution, mice were bled from the retro-orbital sinus. Erythrocytes were eliminated by hypotonic lysis, and the remaining leukocytes were incubated with saturating amounts of phycoerythrin conjugated anti-Ly5.1/CD45.1 antibodies and FITC-labelled anti-Ly5.2/CD45.2 antibodies (Pharminingen, San Diego, CA). Fluorescence was determined using a dual laser FACSCalibur System (Becton Dickinson Immunocytometry Systems, San Jose, CA).

**Linkage analysis:** In this study we used three independent data sets. First, we phenotyped an additional set of 7 newly derived BXD strains 21, and added these data to our database which includes information on 26 BXD strains that we have previously used.
to search for cell cycle QTLs. In addition, we derived two novel datasets consisting of 60 N2 and 43 N6 backcross animals. For linkage analysis we used MapManager QT software (version b28ppc, available at ftp://mcbio.med.buffalo.edu/pub/MapMgr/QT).

We determined the probability of obtaining the observed likelihood ratio statistic (LRS) values by random chance using the permutation test. This test, which is implemented in MapManager QT software, was performed with 2000 permutations at 1 cM intervals (genome-wide for the BXD set, and restricted to chromosome 11 for the N2 and N6 backcross).

**Stem cell purification.** Hematopoietic stem cells were purified using flow cytometry exactly as described before, from pooled marrow cells obtained from 10 B6 and 10 DBA mice. Sorted cells were identified by the absence of a panel of lineage-specific cell surface markers (CD3, CD5, CD8, CD11b/Mac-1, CD45R/B220, and Ly-6G/Gr-1, collectively referred to as Lin-), and the presence of the Sca-1 and c-kit antigen. Viable (propidium iodide-negative) Lin^-Sca-1^c-kit^+ cells were sorted using a dual-laser FACS Vantage instrument (BD Immunocytometry systems, San Jose, CA).

**Subtractive hybridization.** Total RNA was isolated from 10⁵ Lin^-Sca-1^-c-kit^+ B6 and DBA cells using Trizol (Gibco BRL). DNase-treated RNA was reverse transcribed (Superscript II, Gibco BRL) and cDNA was amplified using the SMART PCR cDNA synthesis kit (Gibco BRL). Bi-directional subtractive hybridizations using B6 and DBA amplified stem cell cDNA as “tester” and “driver” reciprocally, were performed with the PCR-Select cDNA subtraction kit (Gibco BRL). All procedures were carried out according to the manufacturer’s suggestions.
Cloning of differentially expressed transcripts. Subtracted cDNA samples were randomly cloned in pCR2-Topo or pCR4-Topo (Invitrogen). A total number of 240 clones (~50% originating from the B6 minus DBA subtraction and 50% derived from the reverse DBA minus B6 subtraction) were sequenced and BLAST searches were performed to identify the transcripts.

Gene array analysis. Amplified, but unsubtracted, B6 and DBA cDNA samples were labelled with $^{33}$P dCTP. Two Mouse Filter Gene Arrays (GF400a, Research Genetics), each containing 5184 cDNA fragment were hybridized with B6 and DBA labelled cDNA samples. Hybridization signals were identified using a Storm phosphor imager system (Molecular Dynamics), and appropriately configured files were analyzed with Pathways Software (Research Genetics). After signal normalization, only spots for which the intensity of either B6 or DBA signal was 3-fold above background were considered. Both in B6 and DBA samples ~12% of the transcripts was found to be expressed in hematopoietic stem cells. We selected the 100 transcripts from both B6 and DBA that showed most pronounced differential hybridization intensities for further study.

Mapping of differentially expressed genes. We tried to determine the chromosomal localization of all (i.e. 440) non-redundant differentially expressed transcripts in silico, using a variety of publicly available databases (predominantly http://www.informatics.jax.org and http://www.ncbi.nlm.nih.gov/UniGene/). We were successful in assigning the chromosomal location to 156 transcripts. Some transcripts were mapped by synteny to their map position in the human genome.

RT-PCR analysis. To independently confirm that transcripts were differentially expressed, we studied the expression pattern of a selected number of chromosome 11-
genes by RT-PCR. To this end purified stem cells or unfractionated bone marrow cells were isolated from an additional cohort of B6 and DBA mice. Also, cells from B6.DBA$^{Scp2}$ congenic mice were used. The primers that were used to amplify differentially expressed transcripts were designed with sequence information as present in Genbank.
RESULTS

Genetic mapping of cell proliferation loci

To map quantitative trait loci associated with variation in cell cycle we performed a genome-wide linkage analysis using cell cycling phenotypes (i.e. fraction of CFU-GM in S-phase) of 33 BXD recombinant inbred strains (analysis of 26 of these strains was previously reported \(^\text{15}\)). We found two markers, D7Mit178 and D11Mit131, that showed suggestive linkage (Table 1) and refer to these QTLs as Stem Cell Proliferation-1 and –2, respectively. For both loci the presence of DBA alleles was associated with higher cell proliferation rates.

Figure 1. Cell cycling phenotypes and genetic analysis of C57BL/6 x B6D2F1 backcross mice. Panel A: The percent progenitors in S-phase for 60 N2 animals is depicted. Black bars refer to animals carrying two B-alleles at D11Mit178, whereas the open bars show heterozygous mice carrying one B- and one D-allele. Panel B: Interval mapping analysis of the N2 generation. The horizontal line indicates the statistical threshold for highly significant linkage, obtained by performing a permutation test with 2000 permutations.
at 1 cM intervals, on chromosome 11 only. Panel C: Interval mapping analysis of the N6 generation.
Panel D: The percentage of CFU-GM in S-phase in wildtype B6 (n=26), DBA (n=29), and in congenic B6.DBA2<sup>Scp2</sup> animals (n=17) is shown.

To confirm linkage a set of 60 (B6 x BDF1)-backcross animals was generated. A distinctive effect of the presence of D-alleles at D11Mit178 (at 47cM) was observed. The phenotype of mice carrying two B-alleles (n=34, mean percentage S-phase cells: 12.4%) was significantly different (p< 2.5 x 10<sup>-9</sup>) compared to mice carrying one B and one D-allele (n=26, mean percentage S-phase cells: 30.2%) (Figure 1A and B).

In contrast to homozygous BXD mice, heterozygous backcross animals provided no evidence for the involvement of Scp1, located on chromosome 7 (data not shown). The backcrossing strategy was continued to derive a congenic mouse strain in which the effect of Scp2 could be investigated in isolation. In N6 mice highly significant linkage was observed with an interval extending from 39 cM to 49 cM on chromosome 11, documenting the strength of this isolated locus (more than 99% of the genome of these animals is of B6 origin) (Figure 1C). Homozygous B6.DBA<sup>Scp2</sup> congenic animals were produced which, as expected, showed a high fraction of cells in S-phase (Figure 1D). The difference between B6 and DBA (p<1.4 x 10<sup>-6</sup>) and between B6 and B6.DBA2<sup>Scp2</sup> (p< 1.2 x 10<sup>-7</sup>) was highly significant. The difference between DBA and B6.DBA2<sup>Scp2</sup> was not significant (p<0.088).

**Stem cell pool size is not affected in B6.DBA<sup>Scp2</sup> congenic mice**

Previous studies by several groups, including ours, have mapped loci regulating stem cell frequency to chromosomes 1, 12, and 18<sup>13,24,25</sup>. These chromosomal locations suggested
that stem cell frequency is maintained by a different set of genes than those affecting cell cycling, mapping to chromosome 11. Our congenic mice offered the possibility to directly assess whether stem cell frequency and proliferation are regulated by distinct genetic pathways. We quantified cobblestone area-forming cell (CAFC) frequencies in bone marrow from B6, DBA, and B6.DBA\textsuperscript{Scp2} mice after 7, 21, and 35 days in culture (Figure 2). As expected, CAFC frequency was substantially higher in DBA compared to B6 bone marrow. However, stem and progenitor cell frequency in the B6.DBA\textsuperscript{Scp2} congenics were identical to B6 values. This demonstrates that Scp2 is not involved in maintaining stem cell pool size.

![Figure 2. Stem cell frequency is not affected by the Scp2 locus on chromosome 11. CAFC assays were performed to quantify the number of progenitors (CAFC day-7, panel A), and more primitive CAFC day-21 and CAFC day-35 cell subsets (panels B and C) in B6, DBA, and B6.DBA\textsuperscript{Scp2} bone marrow cells.]

**B6.DBA\textsuperscript{Scp2} stem cells are defective in long term repopulating ability.**

In embryo- and radiation chimeras in which DBA and B6 stem cells co-exist, highly cycling DBA stem cells, although initially functioning appropriately, cease to contribute to the process of blood cell formation during aging \textsuperscript{26-28}. It has been shown that DBA stem cells can be reactivated after secondary transplantation, demonstrating that DBA
stem cells are present but remain fully dormant in the unmanipulated, aged chimeras\textsuperscript{29}. Conversely, more quiescent B6 stem cells have a major competitive advantage over the long-term, and will eventually outcompete DBA stem cells. Our previous data showing a correlation between cell cycling and mouse lifespan\textsuperscript{12,15}, support a model in which enhanced cellular senescence is induced by high cell turnover rates, and thus would be controlled by $Scp2$. The availability of B6.DBA$^{Scp2}$ congenic mice allowed us to address this question directly by testing the competitive repopulation potential of B6.DBA$^{Scp2}$ stem cells compared to wildtype B6 stem cells. C57BL/6.SJL.Ptpr\textsuperscript{a} (Ly5.1\textsuperscript{+}/CD45.1\textsuperscript{+}) mice were sublethally irradiated and transplanted with 3 x 10\textsuperscript{6} unfractionated B6 or B6.DBA$^{Scp2}$ (Ly5.2\textsuperscript{+}/CD45.2\textsuperscript{+}) bone marrow cells. B6.DBA$^{Scp2}$ stem cells showed a reduced repopulating potential (Figure 3).

Figure 3. Repopulation potential of transplanted stem cells is affected by the $Scp2$ locus on chromosome 11. B6 or B6.DBA$^{Scp2}$ bone marrow cells were intravenously transplanted in sublethally (4Gy) irradiated B6.SJL-CD45.1 congenic recipients. The percentage of donor-derived (CD45.2\textsuperscript{+}) leukocytes was determined by flowcytometry. Data represent mean values +/- 1 SD from two independent experiments. Asterisks indicate significant differences between the two treatment groups ($p<0.05$, T-test).
Since B6 and B6.DBA$^{Scp2}$ bone marrow cells contained an equal number of stem cells (Figure 2), differences in engraftment kinetics must have resulted from qualitative stem cell differences. Although we cannot rule out the possibility that unknown minor histocompatibility antigens may (partially) have induced lower levels of chimerism, no such genes are known to map to the congenic chromosome 11 interval, and therefore we favor the concept that high rates of cell turnover are associated with impaired stem cell functioning.

**Clustering of candidate genes**

By co-transplanting DBA and B6 stem cells into BDF1 recipients we have previously shown that $Scp2$ acts cell-autonomously. Therefore, to detect candidate genes for this locus we identified differentially expressed genes in highly enriched, (and consequently highly identical), Lin$^-$Sca-1$^+$ c-kit$^+$ stem cells isolated by flow cytometry from 10 pooled B6 and DBA bone marrow samples, using bidirectional subtractive hybridization and gene array analysis. After subtractive hybridization 240 B6 and DBA-derived cDNA fragments were cloned and identified by sequence analysis.
Figure 4. Genomic distribution of differentially expressed genes. Transcripts that were shown to be differentially expressed in B6 and DBA purified hematopoietic stem cells using bidirectional subtractive hybridization or gene array analysis were not randomly distributed across the genome, but rather mapped preferentially to chromosome 11.

In parallel, unsubtracted B6 and DBA cDNAs were separately hybridized with DNA arrays. A total of 200 most strongly differentially expressed transcripts were analyzed. We were able to assign the chromosomal localization to 156 transcripts. Interestingly, a disproportionate number of these differentially expressed transcripts was located on chromosome 11 (Figure 4).

A total of 28 genes were shown to map to the averaged sized chromosome 11, which represents ~18% of all differentially expressed transcripts for which a chromosomal map position could be assigned (Table 2). The observed excess of chromosome 11 genes is significantly different from expected random distribution (p<1 x 10^{-4}). In addition, these transcripts were not evenly distributed across chromosome 11 but occurred in three distinct clusters (Table 2, Figure 5).
Figure 5. Clustering of differentially expressed transcripts on chromosome 11. The chromosomal localization of all differentially expressed chromosome 11-transcripts shows that they occurred in three distinct clusters (For exact cM position refer to Table 2). To confirm the differential expression patterns, RT-PCR analyses were performed on a random selection of transcripts using B6, DBA or B6.DBA<sup>Scp2</sup> cDNA. Transcripts are identified by their Unigene (Mm) cluster number.

Assuming a total genome size of ~1500 cM, the distance between the 156 differentially expressed transcript on average is expected to be 156/1500 = ~10 cM. We considered transcripts to be clustered if multiple genes mapped to an interval < 10 cM.

The first cluster centered around 30 cM (6 transcripts were identified that map to this cluster). The largest cluster (containing 11 genes) extended from ~40 to ~50 cM, and coincided exactly with the Scp2 interval (compare Figure 1B and C). The most telomeric cluster covered a region from ~60 to ~78 cM, and contained 9 transcripts.

To confirm the differential expression pattern of transcripts first identified by subtractive hybridization or array analysis, we performed RT-PCR experiments using cDNA obtained from B6, DBA, and B6.DBA<sup>Scp2</sup> bone marrow or stem cells (Figure 5).
When we randomly checked expression of other genes that map closely to those identified by subtractive hybridization or array analysis, we were able to show distinct expression levels in a significant number of them as well (such as \textit{Epx} and \textit{Stat5a}). This indicates that the collection of variably expressed genes within the clusters can be extended.

**Differentially expressed genes map to recombination coldspots**

A more detailed analysis of the \textit{Scp2} cluster showed that differentially expressed transcripts mapped either to a region from 42 to 45 cM, or around 49 cM, whereas no transcripts were found from 45 to 48 cM (Table 2, Figure 5) Also, RT-PCR analysis of randomly selected genes known to map to the silent interval failed to reveal any expression differences (data not shown). Such clustering of differentially expressed genes suggested the involvement of genetic variation affecting higher order chromosomal organization. As most laboratory mice have been derived from a limited number of selected founders there is a number of shared haplotypes across the genome in B6 and DBA mice \textsuperscript{30}. When we assessed the extent of shared haplotypes in the \textit{Scp2} interval (by comparing polymorphism at a series of microsatellite repeats), we found evidence of extensive polymorphism from 38 to 45 cM, and again from 47 to 52 cM, but the silent region from 45 to 47 cM in which we failed to detect differentially expressed genes turned out to be identical in both strains (Figure 6).
Figure 6. Meiotic recombinations, haplotype sharing and differential gene expression in the Scp2 cluster. The frequency of meiotic recombinations within the Scp2 interval was determined in the backcross panel that was used to breed congenic B6.DBA^{Scp2} mice (blue line, left y-axis) and in the BXD recombinant inbred panel (data obtained from http://mickey.umem.edu/MMfiles/MMlist.html, red line, right y-axis). The region of shared haplotypes of B6 and DBA mice (45 to 47 cM) is indicated by the orange box, whereas distinct haplotypes are shown in blue. Differentially expressed stem cell genes (red arrows) clustered in the polymorphic regions in which recombination was severely suppressed.

We made an interesting observation by assessing meiotic recombination frequencies in the Scp2 region using our own backcross panel and the existing BXD panel. In both datasets recombinations in the Scp2 cluster occurred preferentially in the silent region, in which the DBA and B6 genomes are highly homologous. In contrast, recombinations in the bordering regions were strongly suppressed (Figure 6).
DISCUSSION

In this study we demonstrate convergence of a collection of differentially expressed stem cell genes with a locus that is strongly associated with variation in hematopoietic cell turnover between B6 and DBA mice. Differentially expressed transcripts mapped to a polymorphic genomic region that is characterized by suppressed recombination frequencies.

To exclude the possibility that clustering resulted from a preponderance of hematopoietic genes in these regions, we randomly picked 150 genes from the Stem Cell Data Base (http://stemcell.princeton.edu/) which contains a large collection of B6 stem cell specific transcripts. We found no evidence of clustering in this dataset (data not shown), reinforcing the notion that it is the genetic difference between the two strains that causes the clustering. Also, the extent of overrepresentation of chromosome 11 transcripts cannot be attributed to a potentially increased gene density on this chromosome.

Most fundamental to our studies is the issue of how variation in progenitor cell proliferation is induced. Two alternative hypotheses can be provided. First, allelic variation of one of the genes in the Scp2 interval may confer the cell cycling trait. Possible candidates, based upon known functional activity and/or an appropriate (i.e. DBA-like) expression pattern in B6.DBA^{Scp2} congenic mice include Platelet activating factor acetylhydrolase, Ribosomal protein S6 kinase, and Tbox2.

If this “single-gene” hypothesis is correct, then the additional differentially expressed transcripts that are closely linked to the underlying gene are not contributing to the phenotype, and their map position is entirely coincidental. Although we cannot exclude this, we favor an alternative, but more provocative explanation: variation in cell cycle
does not result from a single gene, but rather is the consequence of variable expression of a collection of highly linked genes in stem cells from both parental strains. Genotype-independent phenotypic variability (i.e. reduced penetrance) in the N2, but particularly in the more homogenous N6 generation, would be explained by such a model. The fact that the \textit{Scp2} gene cluster occurs in a genomic interval in which meiotic recombination is suppressed, implies that the cluster is transmitted to the next generation in its entirety. It is therefore tempting to speculate that the strong linkage disequilibrium points to a functional interdependence of genes within the cluster\textsuperscript{35}. Interestingly, several studies have recently documented findings in other model systems that are reminiscent of the data that we show here. First, Caron et al. showed that in the human genome gene expression occurs in highly localized “ridges”\textsuperscript{36}. In addition, suppression of recombination over a large segment has been reported for the cytokine cluster on human chromosome 5q\textsuperscript{37}, a region that has shown strong linkage to various immunological quantitative traits\textsuperscript{38,39}. Intriguingly, the human cytokine cluster is homologous to a segment of mouse chromosome 11 at \textasciitilde 30cM, co-localizing with one of the clusters we report here. Very similarly, an extreme degree of long-range linkage disequilibrium has been reported for a locus on human chromosome 17q21, containing BRCA1\textsuperscript{40}. Again, this segment is homologous to the most telomeric cluster (\textasciitilde 60cM) we report here. Collectively these data indicate that the selective forces that result in effective recombination suppression have been conserved in human and mouse.

It has generally been assumed that the heritable component of quantitative traits result from additive and epistatic interactions of multiple loci, each locus containing a single mutated gene\textsuperscript{16,20}. In this respect it is interesting to note that in our study a number of
DBA-derived genes showed a B6-like expression pattern in the B6.DBA<sup>Scp2</sup> congenic animals (e.g. Zfp147, Mm.4973). This confirms the presence of specific modifying loci in the B6 genome that may prevent the penetrance of a bona fide QTL. Loss of the expected phenotype in congenic mice has repeatedly been observed<sup>41,42</sup>. It has been notoriously difficult to identify genes underlying QTLs<sup>17,43</sup>. Our studies suggest a novel concept which may explain the difficulty of associating individual genes with quantitative traits. Variation at a specific locus may result in altered basal expression of many genes in this region and it is only in the context of the expression pattern of neighbouring genes that biological variation is induced. We speculate that the model we put forward may also apply to certain tumorigenic chromosomal abberations (translocations, deletions, insertions, inversions) in which the altered chromosomal structure affects the expression of a large number of adjacent genes that collectively induce a (subtle) growth advantage.
ACKNOWLEDGEMENTS

The authors thank Anke van den Berg, Geert Harms, Joost Kluiver, and Peter Terpstra for technical assistance and advice.
REFERENCES

1. Lajtha LG. Stem cell concepts. Differentiation. 1979;14:23-34


9. Zsebo KM, Williams DA, Geissler EN, Broudy VC, Martin FH, Atkins HL, Hsu R, Birkett NC, Okino KH, Murdock DC, Jacobsen FW, Langley KE, Smith KA,
Takeishi T, Cattanach BM, Galli SJ, Suggs SV. Stem cell factor is encoded at the Sl locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. Cell. 1990;63:213-224


22. Manly KF, Olson JM. Overview of QTL mapping software and introduction to map manager QT. Mamm Genome. 1999;10:327-334


25


Table 1. Loci most strongly associated with progenitor cell cycling in BXD recombinant inbred mice.

<table>
<thead>
<tr>
<th>Locus</th>
<th>QTL</th>
<th>Chromosome</th>
<th>cM</th>
<th>LRS</th>
<th>%</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>D7Mit178</td>
<td><em>Scp1</em></td>
<td>7</td>
<td>0.5</td>
<td>9.2</td>
<td>22</td>
<td>0.00247</td>
</tr>
<tr>
<td>D11Mit131</td>
<td><em>Scp2</em></td>
<td>11</td>
<td>29</td>
<td>8.4</td>
<td>20</td>
<td>0.00373</td>
</tr>
</tbody>
</table>

A genome-wide search for loci associated with progenitor cell cycling was performed, using the phenotypes of all 33 BXD strains. MapManager QT analysis revealed strongest linkage with these 2 loci. Only linkage with D7Mit178 and D11Mit131 reached LRS values indicative of suggestive linkage. \(^1\)cM indicates the approximate centiMorgan position on the chromosome. \(^2\)Likelihood ratio statistic. \(^3\)% shows the estimated percentage of the variation that is explained by the genotype at that locus. \(^4\)P provides the point-wise p-value.
Table 2. Chromosome 11 genes that are differentially expressed in B6 and DBA hematopoietic stem cells.

<table>
<thead>
<tr>
<th>Gene</th>
<th>UniGene cluster ID</th>
<th>cM position</th>
<th>Expressed more in</th>
<th>Method of detection</th>
<th>Cluster #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ikaros</td>
<td>Mm.12396</td>
<td>6</td>
<td>DBA</td>
<td>SH</td>
<td>-</td>
</tr>
<tr>
<td>Lymphocyte cytosolic protein 2</td>
<td>Mm. 32509</td>
<td>17</td>
<td>DBA</td>
<td>AH</td>
<td>-</td>
</tr>
<tr>
<td>Rad-50</td>
<td>Mm.4888</td>
<td>28.9</td>
<td>DBA</td>
<td>SH</td>
<td>1</td>
</tr>
<tr>
<td>protein phosphatase 2a, alpha isoform EST</td>
<td>Mm.11711</td>
<td>29</td>
<td>DBA</td>
<td>SH</td>
<td>1</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.24328</td>
<td>34</td>
<td>B6</td>
<td>AH</td>
<td>1</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.27282</td>
<td>34,4</td>
<td>B6</td>
<td>SH</td>
<td>1</td>
</tr>
<tr>
<td>Ubiquitin B</td>
<td>Mm.235</td>
<td>35</td>
<td>DBA</td>
<td>SH</td>
<td>1</td>
</tr>
<tr>
<td>EST, Weakly similar to ankyrin B EST</td>
<td>Mm.23940</td>
<td>36</td>
<td>DBA</td>
<td>AH</td>
<td>1</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.27768</td>
<td>42</td>
<td>DBA</td>
<td>SH</td>
<td>2</td>
</tr>
<tr>
<td>Sperm acrosomal protein</td>
<td>Mm.21569</td>
<td>42</td>
<td>B6</td>
<td>AH</td>
<td>2</td>
</tr>
<tr>
<td>Proteasome subunit, beta type 6</td>
<td>Mm.98</td>
<td>43</td>
<td>DBA</td>
<td>SH</td>
<td>2</td>
</tr>
<tr>
<td>Platelet activating factor acetylhydrolase</td>
<td>Mm.56337</td>
<td>44</td>
<td>B6</td>
<td>SH</td>
<td>2</td>
</tr>
<tr>
<td>Homolog of human Gemin4</td>
<td>Hs.19614</td>
<td>44*</td>
<td>B6</td>
<td>SH</td>
<td>2</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.29347</td>
<td>45</td>
<td>B6</td>
<td>AH</td>
<td>2</td>
</tr>
<tr>
<td>Ribosomal protein S6 kinase</td>
<td>Mm.34373</td>
<td>49</td>
<td>DBA</td>
<td>AH and SH</td>
<td>2</td>
</tr>
<tr>
<td>Eosinophil peroxidase</td>
<td>Mm.1315</td>
<td>49</td>
<td>DBA</td>
<td>RT-PCR</td>
<td>2</td>
</tr>
<tr>
<td>Myeloperoxidase</td>
<td>Mm.4668</td>
<td>49</td>
<td>DBA</td>
<td>SH</td>
<td>2</td>
</tr>
<tr>
<td>Zinc finger protein 147</td>
<td>Mm.4973</td>
<td>49</td>
<td>DBA</td>
<td>AH</td>
<td>2</td>
</tr>
<tr>
<td>T-box 2</td>
<td>Mm.4605</td>
<td>49</td>
<td>B6</td>
<td>AH</td>
<td>2</td>
</tr>
<tr>
<td>Hexim1</td>
<td>Mm.8763</td>
<td>56</td>
<td>DBA</td>
<td>AH</td>
<td>3</td>
</tr>
<tr>
<td>STAT 5A</td>
<td>Mm.4697</td>
<td>60,5</td>
<td>DBA</td>
<td>RT-PCR</td>
<td>3</td>
</tr>
<tr>
<td>STAT 5B</td>
<td>Mm.34064</td>
<td>60,5</td>
<td>DBA</td>
<td>SH</td>
<td>3</td>
</tr>
<tr>
<td>Receptor activity modifying protein 2</td>
<td>Mm.1574</td>
<td>61,5</td>
<td>DBA</td>
<td>SH</td>
<td>3</td>
</tr>
<tr>
<td>Homolog of human protein FLJ22561</td>
<td>Hs.180799</td>
<td>65*</td>
<td>B6</td>
<td>SH</td>
<td>3</td>
</tr>
<tr>
<td>Homolog of clathrin, heavy peptide EST</td>
<td>Hs.178710</td>
<td>65*</td>
<td>DBA</td>
<td>SH</td>
<td>3</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.20089</td>
<td>71</td>
<td>B6</td>
<td>AH</td>
<td>3</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.23672</td>
<td>72</td>
<td>DBA</td>
<td>AH</td>
<td>3</td>
</tr>
<tr>
<td>EST</td>
<td>Mm.14076</td>
<td>75</td>
<td>B6</td>
<td>AH</td>
<td>3</td>
</tr>
<tr>
<td>DNA helicase recQ5b</td>
<td>Mm.32869</td>
<td>78</td>
<td>DBA</td>
<td>SH</td>
<td>3</td>
</tr>
</tbody>
</table>

List of chromosome 11 transcripts that are differentially expressed in B6 and DBA stem cells. Three transcripts, indicated by a human Unigene cluster ID and an asterisk, were mapped by synteny to their position in the human genome. AH and SH refer to array hybridization and subtractive hybridization respectively, and indicates the assay in which the transcript was first detected. One transcript, Ribosomal protein S6 kinase, was independently identified by both SH and AH.
A genetic and genomic analysis identifies a cluster of genes associated with hematopoietic cell turnover

Gerald de Haan, Leonid V Bistykh, Ellen Weersing, Bert Dontje, Hartmut Geiger, Natalia Ivanova, Ihor R Lemischka, Edo Vellenga and Gary Van Zant