Data and animal management software for large-scale phenotype screening

Keith A. Ching, Michael P. Cooke, Lisa M. Tarantino, Hilmar Lapp

Genomics Institute, Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, California 92121, USA

Abstract

The mouse *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis program at the Genomics Institute of the Novartis Research Foundation (GNF) uses MouseTRACS to analyze phenotype screens and manage animal husbandry. MouseTRACS is a Web-based laboratory informatics system that electronically records and organizes mouse colony operations, prints cage cards, tracks inventory, manages requests, and reports Institutional Animal Care and Use Committee (IACUC) protocol usage. For efficient phenotype screening, MouseTRACS identifies mutants, visualizes data, and maps mutations. It displays and integrates phenotype and genotype data using likelihood odds ratio (LOD) plots of genetic linkage between genotype and phenotype. More detailed mapping intervals show individual single nucleotide polymorphism (SNP) markers in the context of phenotype. In addition, dynamically generated pedigree diagrams and inventory reports linked to screening results summarize the inheritance pattern and the degree of penetrance. MouseTRACS displays screening data in tables and uses standard charts such as box plots, histograms, scatter plots, and customized charts looking at clustered mice or cross pedigree comparisons. In summary, MouseTRACS enables the efficient screening, analysis, and management of thousands of animals to find mutant mice and identify novel gene functions. MouseTRACS is available under an open source license at http://mousetracs.sourceforge.net.

Correspondence to: Keith A. Ching; E-mail: mousetracs@users.sourceforge.net

Introduction

Drug discovery methods using the *N*-ethyl-*N*-nitrosourea (ENU) mutagenesis of mice (Russ et al. 2002) aim to reveal novel gene functions by associating phenotypic changes with genes altered via mutation. The availability of economical large-scale sequencing, single nucleotide polymorphism (SNP) detection, and genomic SNP maps from multiple strains of mice (Pletcher et al. 2004) have enabled researchers to locate the genomic positions of disease-causing mutations in mice. This premise underlies the strategy and promise for mouse ENU mutagenesis programs. ENU is a chemical that causes random genome-wide point mutations (Russ et al. 2002). Many of the changes caused by these mutations will be unnoticeable; however, by chance a mutation may occur in a gene that produces a phenotypic change. These physical, measurable changes in phenotype can be revealed by screens designed to identify mice with abnormal characteristics such as high cholesterol levels, deficient immune function, or altered behavior. Once such a "pheno-deviant" mouse is identified, the gene responsible can be identified by mapping the mutation to the genomic sequence via outcrossing breeding strategies, SNP sequencing, and positional cloning (Nelms and Goodnow 2001). However, the cost and labor of screening, breeding, and analyzing the large numbers of animals required to produce, identify, and map mutants make automated information management a necessity.

In this article we present MouseTRACS, an informatics solution for mouse data and animal management to increase vivarium efficiency, analyze screening data, and aid in positional cloning. MouseTRACS currently handles data from approximately 9000 to 10,000 mice annually in over 20 different screens and data from an animal population
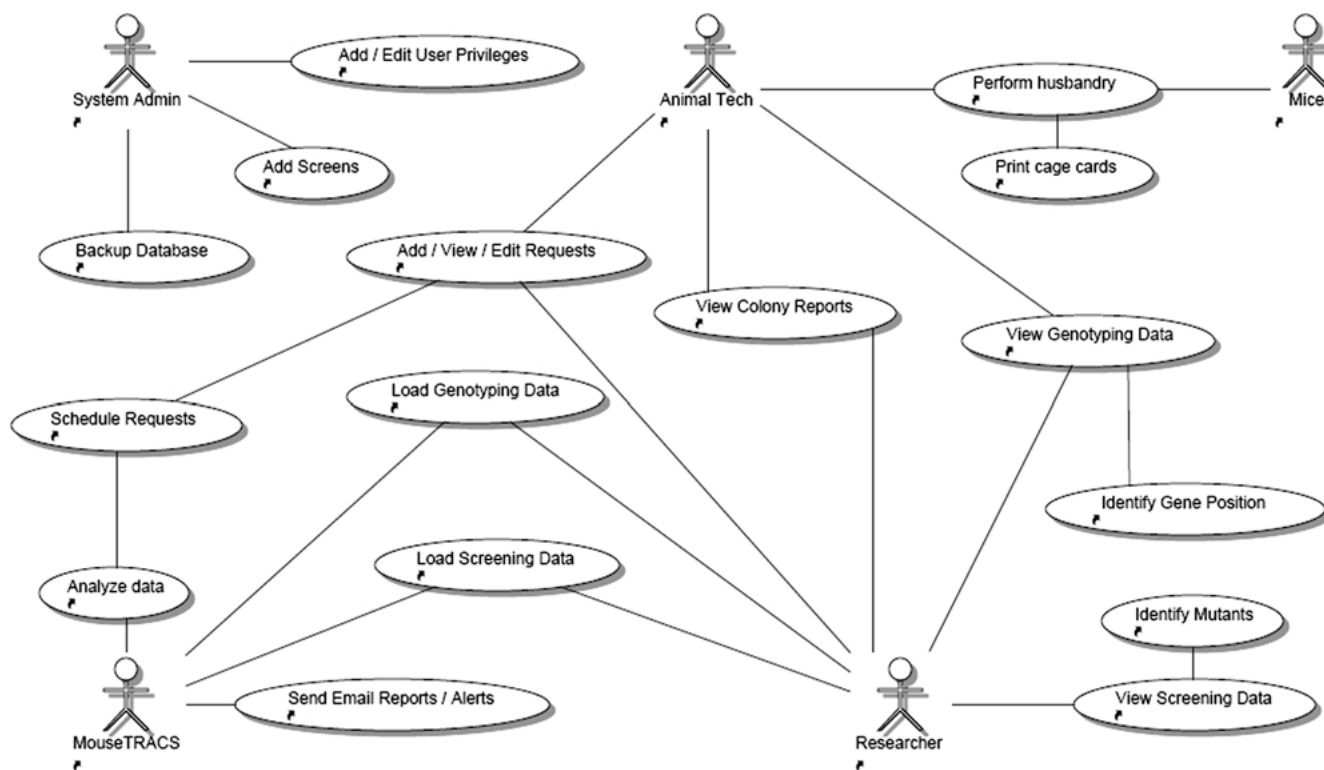
Fig. 1. A use case diagram outlines the main functionality and value of MouseTRACS. Users (stick figures) interact with the system to accomplish specific tasks (ovals). Animal technicians perform animal husbandry and also view animal management requests made by researchers. Researchers load and view the screening data. The MouseTRACS system runs automated analysis scripts that flag data and make animal management requests. The system administrator maintains the database by backing up data, adding users, and adding new phenotyping screens.

of over 20,000 for breeding, inheritance testing, and mutation mapping. It has been used to manage animals, store genotypes, and/or flag screening data for a number of successful cloning projects such as inositol (1,4,5) trisphosphate 3 kinase B (ITPKB) (Wen et al. 2004), c-Myb (Sandberg et al. 2005), aquaporin-2 (Lloyd et al. 2005), and over 30 other strains currently under investigation.

MouseTRACS was implemented using Perl (http://www.perl.com), Java (http://java.sun.com), and MySQL (http://www.mysql.com). Although MouseTRACS was specifically designed for managing GNF's in-house mouse ENU screening program, the mouse colony management system is also configurable as a stand-alone solution for breeding large numbers of animals. It may find utility in small companies, universities, or other institutes with animal facilities in need of an economical yet capable animal management informatics solution.

Overview

The main functionality of MouseTRACS can be diagrammed as a series of use cases from the user's point of view. As shown in Fig. 1, there are four

main groups of users: the system administrator, animal technicians, researchers, and programmatic scripts. The system administrator's job is to backup the data, modify user privileges, and add new data items like users, screens, protocols, and mouse backgrounds. In contrast, animal technicians perform animal husbandry, view requested tasks, and examine inventory. Researchers, on the other hand, focus on loading, viewing, and interpreting the screening and genotyping data. Based upon the strength of the data, they decide what actions to perform on the animals. These actions are immediately conveyed to the animal technicians in the form of requests entered into the system. The remaining "user" is the system itself which can programmatically access MouseTRACS functions. The system must load the data and perform routine data analysis. Based upon the analysis results, the system also places automated requests and sends out email alerts for important events or findings.

Software implementation

MouseTRACS was largely written in Perl with Java applets. MySQL was used as the relational database

management system (RDBMS). The Web applications run on Unix-based systems such as Linux and Mac OSX and use the Apache Web server (http://www.apache.org) (see supplementary materials for details).

Screening data

Data from phenotype screens is loaded by parsing data files of varying formats such as tab-separated text, comma-separated text, binary, and machine output into a standard format. A validator script checks the dates, mouse numbers, and screening data for mistakes. Once the file passes the validator script, it is automatically loaded into the database at scheduled intervals.

MouseTRACS will automatically flag the test outliers and reschedule the corresponding mice for retesting in the appropriate assay. Or, if the new data is retest data, it will be compared with previous data and marked as either confirming or nonconfirming. After data loading and analysis, completed retest requests are automatically taken off the task list. Emails are sent to researchers when mutant family lines accumulate multiple affected animals, i.e., a single mutagenized founder sires many phenotypically remarkable progeny (see supplementary material for details).

Data viewer

Mutant identification begins with finding mice that present an aberrant phenotype due to mutated, homozygous recessive alleles in the G3 population (Nelms and Goodnow 2001). In the data viewer, outlier phenotypes are highlighted to reveal multiple affected individuals that are derived from the same founder line. For example, Fig. 2 shows that mutants from founder line No. 7 are readily apparent (mice 206, 283, 285) in the data viewer. An expected 1/8 proportion of mutants should be evident if the alleles follow Mendelian inheritance (Nelms and Goodnow 2001).

The data viewer also serves as the interface for researchers to make requests on mice. Requests on animals traditionally come in the form of informal emails to the responsible technician. Prioritizing, organizing, and tracking the deluge of requests can sap much of a technician's time with constant requests and confirmations of task completion. Request tracking alleviates this issue by providing a task list of outstanding requests which is automatically organized for the technician. An animal technician can view the "to do" lists of the requests that are automatically checked off as they are completed.

Researchers can add to the list and check on the status of their requests without bothering the technician. Importantly, an audit trail is provided to track request details, when the request was made, and when it was completed.

Affected animals can be kept for further study by selecting "save" in a pull-down menu on the row of the desired mouse. Other actions on mice such as kill, breed, retest, in vitro fertilization, or genotyping can also be requested. When multiple affected animals from the same founder line are saved, they are bred to test for inheritance of the mutant phenotype. If the phenotype is heritable, then a mapping cross (Nelms and Goodnow 2001) can be requested and resulting animals can be scheduled for genotyping. Otherwise, all animals from a nonheritable line can be scheduled for termination by marking the line as "retired."

MouseTRACS generates graphs by selecting a subset of data by test, date, genetic background, generation, sex, and/or allele on a Web form. Available chart types include box plots, scatter plots, histograms, and dendrograms for standard clustering algorithms. Chart rendering uses the open source "R project for statistical computing" (R Development Core Team 2004). A Perl script programs R to import the data and generate the chart. Raw data can be exported into tab-separated files for import into sophisticated visual tools like Spotfire (http://www.spotfire.com/).

A number of predefined charts are available to help identify individual mutant mice and mutant founder lines in the initial G3 screen and inheritance crosses. Interpedigree charts show the range of values for a given test in each pedigree and help to delineate the expected range for normal test values. Figure 3A shows that a group of G3 animals from founder line No. 7 have low numbers of cells expressing CD3 (a T-cell lymphocyte marker) (Kane et al. 2000), while a group of animals derived from founder No. 1 have high CD3 cell counts. Intrapedigree clustering attempts to separate the mutants from the unaffected animals by using all of the tests in a given assay. For example, B220, CD3, CD4, and CD8 are used for the flow cytometry screen. As shown in Fig. 3B, the Partitioning Around Medoids (PAM) clustering method clusters the G3 mice from founder No. 7 and plots them according to the two tests that account for most of the variability (Kaufman and Rousseeuw 1990). These clusters can also be visualized by a simple HTML heatmap. Other common clustering algorithms from the R cluster package (Rousseeuw et al. 2004) are available such as agnes (Agglomerative Hierarchical Clustering), diana (Divisive Hierarchical Clustering), or fanny (Fuzzy Analysis Clustering).
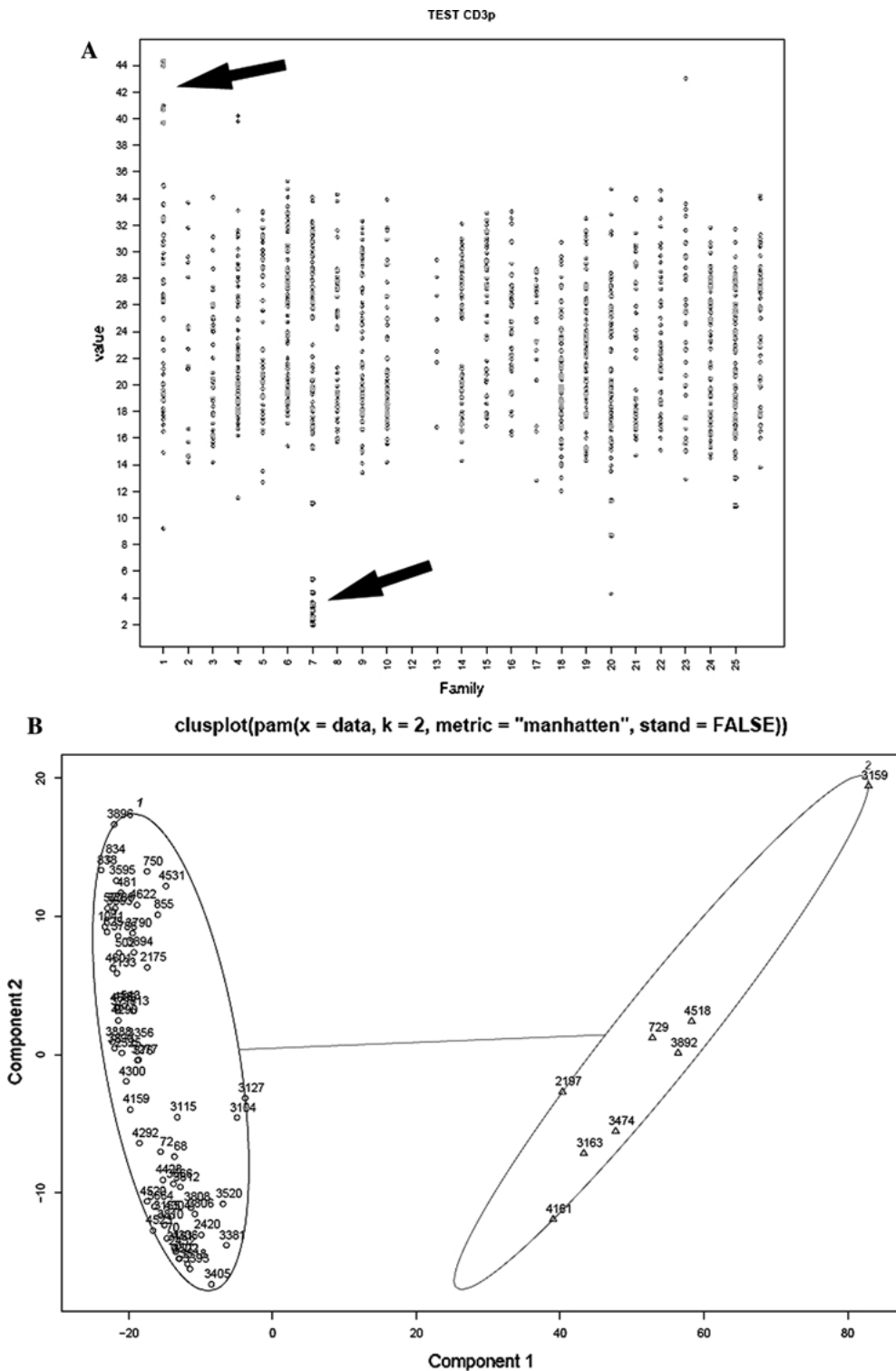
Fig. 2. Identification of founder No. 7. This sample screenshot of the data viewer shows G3 mice from founder line No. 7 that were identified as low B-cell and low T-cell mutants. Mice are arranged in rows with summary information such as sex, birth date, generation, background, allele, and current requests. Screening data are shown in columns with the test name, date performed, and test value. To highlight outliers, flagged data is colored according to *z*-score where highs are a shade of red and lows a hue of blue. Because of inherent noise in the screening data, mice that flag for a test are autoscheduled for a confirming retest. Confirming test results are boxed with a yellow outline and nonconfirming tests are boxed in green. Tests are organized into views of logically related assays such as flow cytometry and hematology. Data can be searched by founder, background, allele, generation, date of test, mouse ID, investigator, and various combinations of the above.

Colony management

MouseTRACS provides a stand-alone animal management system that models the workflow of the animal technicians. Technicians create a virtual cage and add virtual mice via a Web form. Mice are automatically assigned a unique identifier by the database. Dropdown menus allow for setting attributes such as genetic background, alleles, generation, genotype, phenotype nicknames, IACUC protocol numbers, investigators, and comments.

An animal technician sets up a breeding cage and fills out a breeding card form like the cage card form. When a litter is born, the birth date and the total number of pups are entered. Upon weaning of the pups from the mother, the technician enters the weaning date and the total number of surviving males and females. The computer automatically creates the mice within the system according to the attributes on the breeding card, distributes them into cages by sex, and prints the corresponding cage cards. Technicians record subsequent litters by clicking a button to add

Fig. 3. (**A**) Interpedigree graph compares test values between pedigrees. Mutants from founder family No. 7 can be seen as a cluster of low CD3 values relative to test results from other founder lines. Founder family No. 1 mutants with high values can also be seen. Test values for CD3 are plotted on the y axis while founder line is plotted on the x axis. (**B**) Intrapedigree clustering compares test values within a pedigree. Test results for all tests in the flow cytometry assay for G3 mice from founder line No. 7 were clustered using the PAM (Kaufman and Rousseeuw 1990) clustering method. The top two principal components are used to plot the mouse ID on the x and y axes. Mice are separated into k clusters, two in this case. Mutants are represented by the right ellipse while unaffected mice reside in the left ellipse. The line connecting the center of the ellipses represents the distance between clusters. Overlapping ellipses would indicate poor cluster separation due to no mutants or more than two populations of mice.

another entry to the breeding card. The computer creates a new breeding card using the information from the previous breeding card. Technicians can print updated breeding cards to reflect new litters.

The printed cage cards contain important reference information. As shown in Fig. 4A, the card displays the sex, generation, genetic background, alleles, birth date, genotypes, mouse ID, and number

A

59991	59991	2005-08-05	Gekakis	02-32

Mother 153092   Father 130448
C57BL/6J-   B6.ENU-217(9019)

Strain B6.ENU-217(9019)	F5N1	2005-04-30

grp	sex	tag	val	ID	Gtyp
0	M	172449	.	172449	1+/+
0	F	172454	.	172454	1+/-
0	F	172455	.	172455	1+/-

PEESTERS
PEESTERS
PEESTERS

B

59991	2005-08-05	Gekakis	02-32

M.mother C57BL/6J-	Mother B6.ENU-217(9019)	F5N1
M.father B6.ENU-217(9019)	2005-04-30 172454 172455	. 0
F.mother C57BL/6J-	Father B6.ENU-217(9019)	F5N1
F.father B6.ENU-217(9019)	2005-04-30 172449	. 0

2005-08-05	B6.ENU-217(9019)	F6N1

#	P	W	DOB	T:F:M	Comment
1	00-00-00	05-09-14	05-08-24	7:5:2	.
2					
3					
4					
5					
6					
7					
8					
9					
10					

Fig. 4. (**A**) Sample of a printed weaning card shows the cage number, parental cage, cage creation date, protocol, protocol number, parental information, and information about the mice in the cage. (**B**) A printed breeding card shows parental information for the breeding mice as well as litter information.

of animals in the cage. In addition, the card shows the protocol number, cage ID, setup date, and responsible investigator as well as the parental genetic background and parental mouse IDs. Any comments on the mice made in the database are also printed on the card.

The breeding card in Fig. 4B shows all the information required for breeding such as the genetic lineage of the father's parents and the mother's parents. Previous and current litters are printed with the wean dates, dates of birth, and number born and weaned along with any comments from the technicians. Typical comments include observations like found dead, missing, eaten, or other behavioral observations that could be important. For instance, the aquaporin-2 mutants described by Lloyd et al. (2005) were initially noted for their excessive urination and water consumption.

Inventory control and planning

Much of the cost savings and productivity benefits of MouseTRACS comes from inventory reporting and control. Researchers outside the vivarium can inventory and track mice without consulting the animal technicians on a daily basis. Dynamically generated inventory reports detail the numbers of mice broken down by investigator, genetic background, generation, cage, and facility location. The report also counts how many mice are in each category of alive, dead, born, weaned, or tested. In the reports, mice are Web-linked to cage information and test results.

As vivarium space becomes limited, it is important to clear unnecessary cages efficiently. Large groups of animals easily can be scheduled for retirement. Many database reports were built to identify infertile or nonmutated mice and to limit breeding to the minimum required. Thus, mice with specific genetic backgrounds and generations can be restricted to a certain number of animals produced such that new breeding and weaning operations are blocked and the pertinent investigators are alerted by autogenerated emails. From an animal welfare perspective, electronically tracking and limiting the breeding of animals minimizes unnecessary animal use and extended time on the shelf. Furthermore, this helps to keep down costs and to use the available space effectively for as many studies as possible.

Regulatory compliance and auditing

Institutional Animal Care and Use Committee (IACUC) Protocol tracking is an important component of regulatory compliance with laws regarding laboratory animals used for research (http://www.iacuc.org). MouseTRACS automatically assigns IACUC protocol numbers to offspring and can generate reports of animal usage to help satisfy compliance and auditing requirements.

Pedigree documentation and visualization

Complex breeding schemes can be difficult to track and manage. Record-keeping for breeding multiple
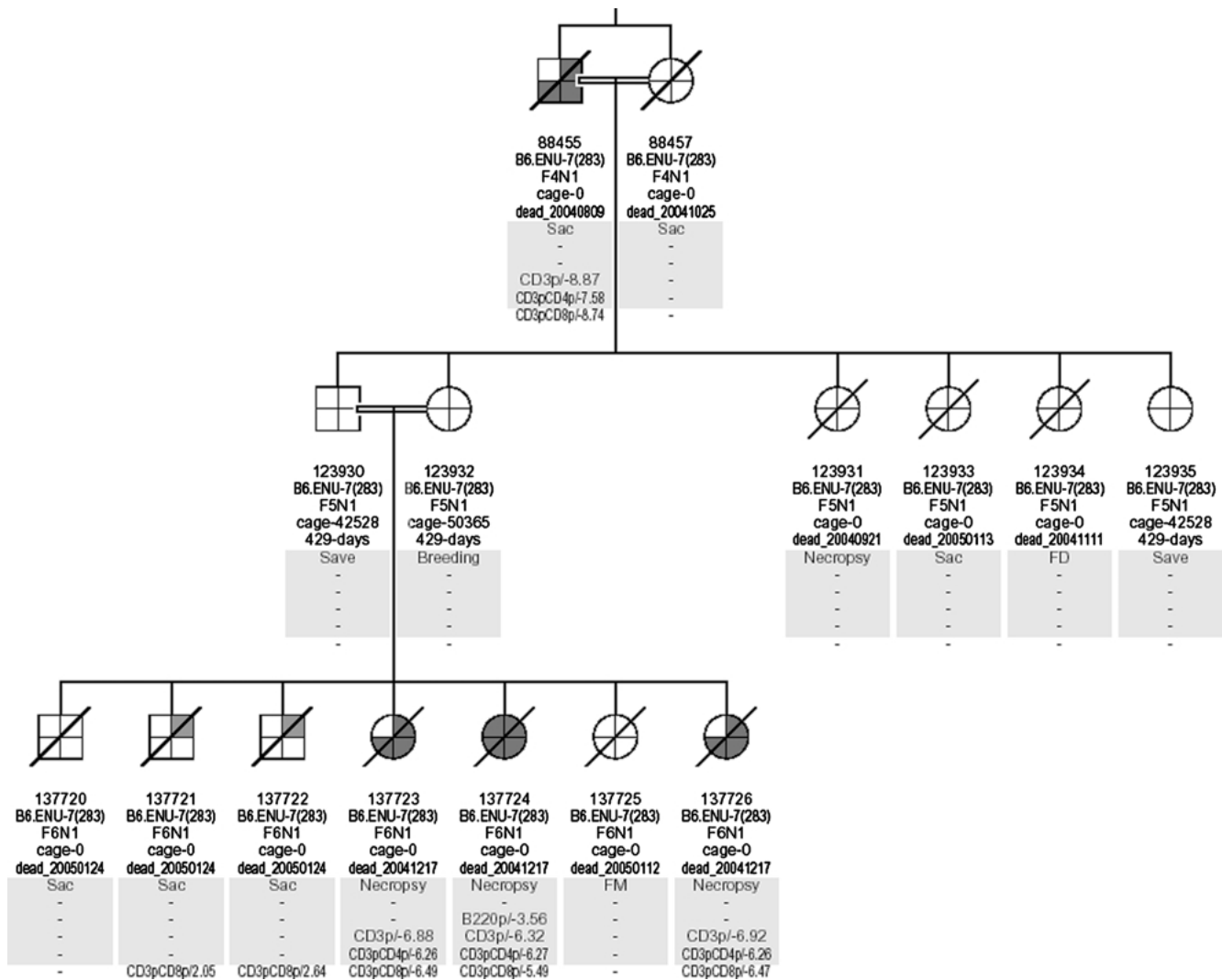
Fig. 5. Pedigree from a founder No. 7 mating with overlaid data. Double lines between parents denote a sibling intercross. Slashes indicate deceased animals. Light and dark shaded sections show traits that are higher and lower than usual, respectively. Test names and *z*-scores are shown for flagged test values.

lines is time-consuming, tedious, and error-prone. MouseTRACS allows for the visualization of the lineage information stored in the database by using a slightly modified version (see supplementary materials) of the open source Madeline v0.933 genetic linkage software (Trager 2001). The pedigree view integrates data from the phenotyping screens and lineage information in order to examine patterns of inheritance. As shown in Fig. 5 mutants from founder No. 7 can be identified readily by color and listed *z*-scores to examine the proportion of mutants to wild-type and intermediate phenotypes.

Integration with genotyping data

MouseTRACS stores and analyzes the genotyping data that is used to map mutant genes or identify the genetic lineage of mutant crosses. Mapping data is transferred from an Oracle database that stores the results obtained from the Sequenom MassARRAY® System (http://www.sequenom.com) into a SNP database implemented in MySQL. Because vendor changes to the database occur periodically and occasionally different vendors are used, MouseTR-ACS uses its own separate schema. Using a neutral format maintains vendor independence. Users also genotype mice via PCR or other direct methods. These data are imported into the same SNP database via a tab-delimited text import in the same fashion as the screening data import.

The data viewer displays genotype information alongside the allele information. Clicking on the genotype will show the details of the call such as nucleotide alleles, date, quality of the call, and
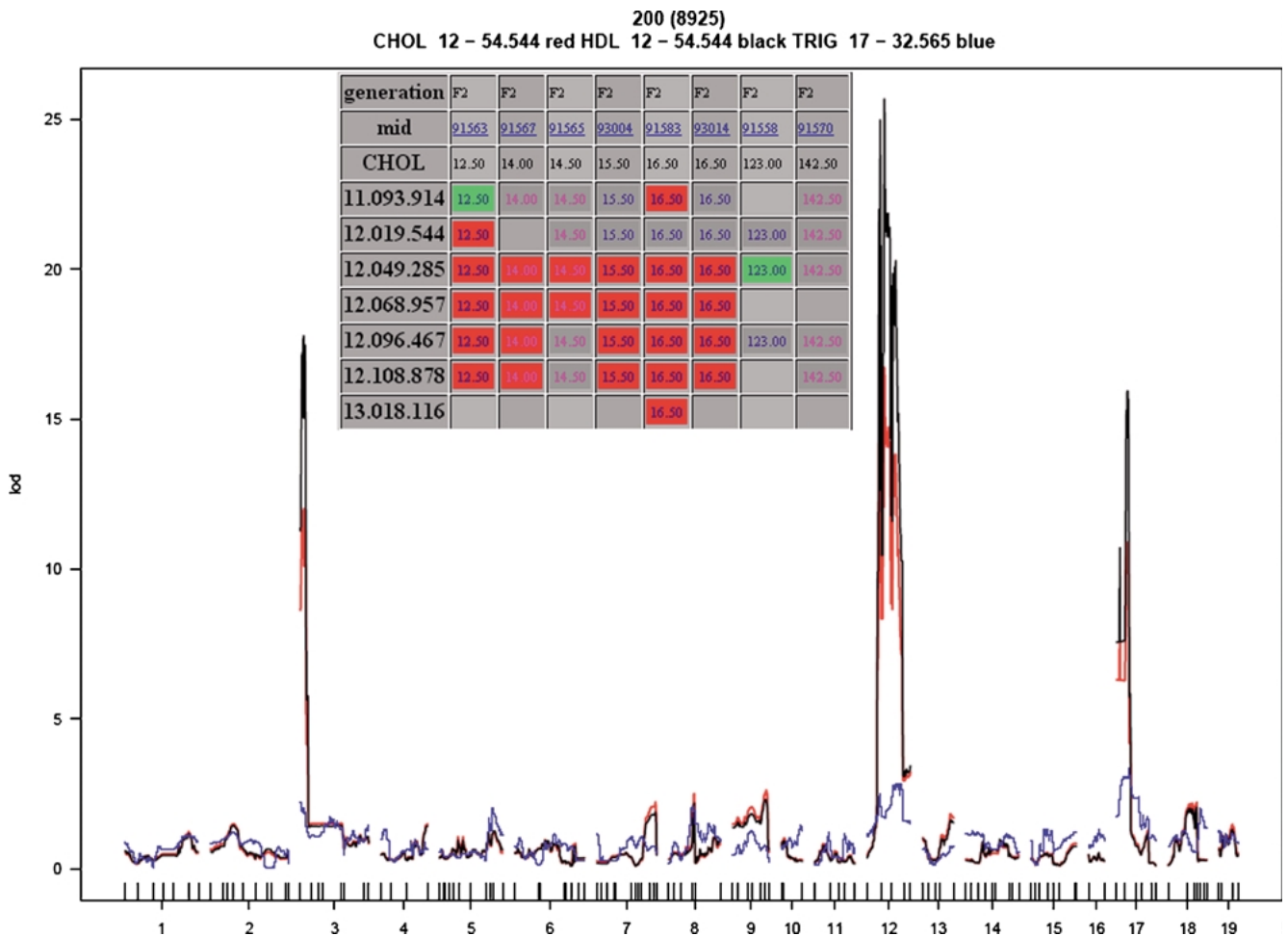
Fig. 6. Automated QTL mapping. The inset shows a Web-generated SNP interval map that displays test results in the context of genetic background. The genomic location is listed in the first column as chromosome.megabases.kilobases. Results for males are blue while females are pink. Green cells indicate the mapping strain. Red cells show the mutant strain while gray cells show heterozygous SNP calls. Using R and the qtl package, the generated LOD plot shows genomic locations that are correlated with measured phenotypes. An initial genome scan of eight mice for two affected phenotypes (red and black) show four major peaks. One unaffected phenotype of triglycerides (blue) is used as a negative control. Peaks that are high in cholesterol (red) and HDL (black) but not triglycerides are potential loci for genes with causative mutations.

operating technician. The genotype information is also displayed next to the allele in the colony view, printed on the cage cards, and is included when exporting the screening data. Some plots will graph data based on the genotype of the mouse.

Automated quantitative trait loci (QTL) mapping is performed using the R qtl package (Broman et al. 2003). (http://www.biostat.jhsph.edu/~kbroman/qtl/) To map SNP nucleotide calls to either the background strain or the mapping strain, roughly 450,000 SNPs from 48 strains of inbred mice (Pletcher et al. 2004) were loaded into the SNP database. Perl scripts then format the mapping information into a ''map.-cross'' file suitable for import into the R qtl package. The ''scanone'' imputation method (Broman et al. 2003) is used to plot up to three phenotypes on a LOD

plot. The LOD plot indicates regions of the genome where SNP markers of a particular genotype correlate with the phenotype in question. Figure 6 shows that cholesterol and high-density lipoprotein (HDL) levels (red and black) highly correlate with genotype on Chromosome 12 but that triglyceride levels (blue) do not. The corresponding genomic interval can be visualized in detail via an HTML table as seen in the inset of Fig. 6. At the single-mouse level, low cholesterol levels correlate with SNP markers from the mutant strain (red) in a haplotype block. In contrast, animals with normal cholesterol are genotyped as heterozygous (gray) or from the mapping strain (green). This information can be used as a starting point for interval selection and refined mapping by using dense SNP markers in the putative region.

Discussion

Although MouseTRACS was developed independently, it shares much functionally, such as an animal management system, experimental data storage, and data analysis, with other mouse ENU informatics solutions that have been published since its conception. We discuss how MouseTRACS implements these features compared with the following four ENU informatics systems published over the last five years: Mutabase from Medical Research Council Harwell (MRC) (Strivens et al. 2000), MouseNet from Forschung für die Gesundheit (GSF) (Pargent et al. 2000), MuTrack from The Tennessee Mouse Genome Consortium (TMGC) (Baker et al. 2004), and MUSDB from RIKEN GSC (Masuya et al. 2004).

As pointed out by Baker et al. (2004), engineering a generic informatics solution is difficult. Each ENU program has different screening workflows, novel data, and various methods of animal management that require customization. The applicability, interchangeability, and usability of any system for another ENU program is limited by how generalized the software was written, the availability of the source code, the software implementation language, and the RDBMS. Because of the extensive customization, different ENU informatics systems are not interchangeable, yet they must solve similar problems.

Two approaches to data loading have been used. MouseTRACS, MuTrack, and MouseNet parse files and load them from a centralized location. In contrast, the approach taken by MUSDB uses direct data transfer from analytical devices to the database. Direct transfer eliminates the need for transferring files but requires the development of custom client software for each device.

Automated screening data analysis is performed by MouseTRACS and MuTrack to flag outliers. Mutabase and MUSDB provide on-demand statistical calculations. The advantage to offline statistics calculations is that the results are available to everyone and can be stored and compared over time. For instance, MouseTRACS provides graphs to examine trends in flagging thresholds which can help identify problems with instrumentation or reagents. However, on-demand statistics provide greater flexibility and control in defining the data set and specific methods used. For these cases, MouseTRACS provides a data export capability for researchers who want to perform their own, customized statistics.

Tracking the breeding operations and requests on thousands of animals is the main logistical challenge for a large animal facility. All ENU informatics systems provide animal husbandry functionality to address these requirements. By necessity, the pedigree information must be carefully preserved or else mutation mapping will become impossibly convoluted. Printing bar-coded animal management cards greatly enhances productivity because hundreds of cards are printed every day for the thousands of cages.

The animal husbandry functionality of MouseTRACS can be decoupled from the ENU screening functions via a configuration file. This enables stand-alone use as animal management software. In this manner, parts of the ENU functionality such as genotype tracking can be turned on if the need later arises. The breeding, weaning, tracking, and cage card printing functionality alone can provide much value over a pen and paper operation with minimal cost. For instance, the use of MouseTRACS as an animal management system for GNF's Pharmacology Animal Research facility helps technicians maintain productivity with rapidly increasing numbers of animals. We hope that MouseTRACS can provide similar benefits to other facilities.

Compared with other ENU informatics systems, genotype tracking and integration is the most distinguishing feature of MouseTRACS. Users can quickly evaluate mice and mutants because MouseTRACS provides convenient access to both genotype and phenotype in an integrated fashion. Genotyping information was previously stored on personal computers in Excel spreadsheets and in an inaccessible, proprietary RDBMS schema that could be deciphered only by the MassARRAY® technicians. Furthermore, researchers had to match genotypes to the screening data by hand. They had to wait for hand-generated genomic interval maps that could not immediately reflect new genotyping calls, phenotyping data, or additional animals. To address these bottlenecks, MouseTRACS dynamically places genotype calls and genomic position in context with phenotyping data and allows for automated QTL mapping. Maps can be generated on demand by anyone and will always reflect the latest data available in the database.

In summary, MouseTRACS is a configurable animal management system that enables the tracking and management of hundreds of thousands of animals from birth to death. Without an informatics solution, animal management would be a costly, tedious, and error-prone deluge of paper records, email requests, and Excel spreadsheets. MouseTRACS provides the benefits of electronic records management, experimental data access and analysis,

regulatory compliance, and inventory cost control. The additional advantages of low hardware requirements, flexible configurability, and freely modifiable code provide compelling reasons to use MouseTR-ACS as a low-cost, full-featured animal information management solution.

Acknowledgments

References

1. Baker E, Galloway L, Jackson B, Schmoyer D, Snoddy J (2004) MuTrack: a genome analysis system for large-scale mutagenesis in the mouse. BMC Bioinformatics 5, 11

2. Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889–890

3. Kane LP, Lin J, Weiss A (2000) Signal transduction by the TCR for antigen. Curr Opin Immunol 12(3), 242–249

4. Kaufman L, Rousseeuw PJ (1990) *Finding Groups in Data: An Introduction to Cluster Analysis* (New York: Wiley)

5. Lloyd DJ, Hall FW, Tarantino LM, Gekakis N (2005) Diabetes insipidus in mice with a mutation in aquaporin-2. PLoS Genet 1(2), e20

6. Masuya H, Nakai Y, Motegi H, Niinaya N, Kida Y, et al. (2004) Development and implementation of a database system to manage a large-scale mouse ENU-mutagenesis program. Mamm Genome 15(5), 404–411

7. Nelms KA, Goodnow CC (2001) Genome-wide ENU mutagenesis to reveal immune regulators. Immunity 15(3), 409–418

8. Pargent W, Heffner S, Schable KF, Soewarto D, Fuchs H, et al. (2000) MouseNet database: digital management of a large-scale mutagenesis project. Mamm Genome 11(7), 590–593

9. Pletcher MT, McClurg P, Batalov S, Su AI, Barnes SW, et al. (2004) Use of a dense single nucleotide polymorphism map for in silico mapping in the mouse. PLoS Biol 2(12), e393

10. R Development Core Team (2004) R: *A language and environment for statistical computing* (Vienna: R Foundation for Statistical Computing), ISBN 3-900051-07-0. Available at http://www.R-project.org

11. Rousseeuw P, Struyf A, Hubert M, Hornik K, Maechler M (2004) cluster: Functions for clustering. R package v1.9.6

12. Russ A, Stumm G, Augustin M, Sedlmeier R, Wattler S, et al. (2002) Random mutagenesis in the mouse as a tool in drug discovery. Drug Discov Today 7(23), 1175–1183

13. Sandberg ML, Sutton SE, Pletcher MT, Wiltshire T, Tarantino LM, et al. (2005) c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation. Dev Cell 8(2), 153–166

14. Strivens MA, Selley RL, Greenaway SJ, Hewitt M, Liu X, et al. (2000) Informatics for mutagenesis: the design of mutabase–a distributed data recording system for animal husbandry, mutagenesis, and phenotypic analysis. Mamm Genome 11(7), 577–583

15. Trager EH (2001) Open source Madeline v0.933 pedigree drawing program. Available at http://eyegene.ophthy.med.umich.edu/madeline-0.933/index.html

16. Wen BG, Pletcher MT, Warashina M, Choe SH, Ziaee N, et al. (2004) Inositol (1,4,5) trisphosphate 3 kinase B controls positive selection of T cells and modulates Erk activity. Proc Natl Acad Sci USA 101(15), 5604–5609