

was generated and custom software (pPicker; Perlegen Sciences, Inc., Mountain View, CA) was designed to choose a minimal set of nonredundant primers that yield maximum coverage of chromosome 21 sequence with a minimal overlap between adjacent amplicons. LR-PCR reactions were performed using the Expand Long Template PCR Kit (Roche Biosciences, Palo Alto, CA) with minor modifications.

13. LR-PCR targets were prepared as previously described with some modifications (7). For each wafer hybridization, corresponding LR-PCR products were pooled and purified using Qiagen tip 500 (Qiagen, Valencia, CA). A total of 280 µg of purified DNA was fragmented using 37 µl of 10× One-Phor-All buffer PLUS (Promega, Madison, WI) and 1 unit of DNAase (Life Technologies/Invitrogen, Carlsbad, CA) in 370 µl total volume at 37°C for 10 min, which was then heat-inactivated at 99°C for 10 min. The fragmented products were end labeled using 500 units of Tdt (Boehringer Mannheim) and 20 nmoles of biotin-N6-ddATP (DuPont NEN, Boston, MA) at 37°C for 90 min and heat inactivated at 95°C for 10 min. The labeled samples were hybridized to the wafers in 10 mM tris-HCl (pH 8), 3 M tetramethylammonium chloride, 0.01% Tx-100, 10 µg/ml denatured herring sperm DNA in a total volume of 14 ml per wafer at 50°C for 14 to 16 hours. The wafers were rinsed briefly in 4× SSPE, washed three times in 6× SSPE for 10 min each, and stained with streptavidin R-phycoerythrin (SAPE; 5 ng/ml) at room temperature for 10 min. The signal was amplified by staining with an antibody against streptavidin (1.25 ng/ml) and by repeating the staining step with SAPE. The wafers were scanned using a custom-built confocal scanner.
14. A combination of previously described algorithms (7), was used to detect SNPs based on altered hybridization patterns.
15. Consistent failure of LR-PCR in all samples analyzed accounts for 15% of the 35% false negative rate. The remaining 20% false negatives are distributed between bases that never yield high-quality data (10%) and bases that yield high-quality data in only a fraction of the 20 chromosomes analyzed (10%). In general, it is the sequence context of a base that dictates whether or not it will yield high-quality data. Our finding that approximately 20% of all bases give consistently poor data is very similar to the finding that approximately 30% of bases in single dideoxy sequencing reads of 500 bases have quality scores too low for reliable SNP detection [D. Altschuler et al., *Nature* 407, 513 (2000)]. The power to discover rare SNPs as compared to more frequent SNPs is disproportionately reduced in cases where only a limited number of the samples analyzed yield high-quality data for a given base. As a result, our SNP discovery is biased in favor of common SNPs.
16. D. L. Hartl, A. G. Clark, *Principles of Population Genetics* (Sinauer, Sunderland, MA, 1997), pp. 57–60.
17. The International SNP Map Working Group, *Nature* 409, 928 (2001). We compared the overlap of 15,549 chromosome 21 SNPs discovered by The SNP Consortium (TSC) with the SNPs found in this study. Of the TSC SNPs, 5087 were found to be in repeated DNA and were not tiled on our wafers. Of the remaining 10462 TSC SNPs, we identified 4705 (45%).
18. J. C. Stephens et al., *Science* 293, 489 (2001)
19. Y. Fu, W.-H. Li, *Genetics* 133, 693 (1993).
20. G. Marth et al., *Nature Genet.* 27, 371 (2001).
21. Z. Yang et al., *Nature Genet.* 26, 13 (2000).
22. In the course of SNP discovery, we identified 339 SNPs that appeared to have more than two alleles. These SNPs were not included in any analyses.
23. L. Excoffier, M. Slatkin, *Mol. Biol. Evol.* 12, 921 (1995).
24. S. Michalatos-Beloin et al., *Nucleic Acids Res.* 24, 4841 (1996).
25. S. E. Hodge, M. Boehnke, M. A. Spence, *Nature Genet.* 21, 360 (1999).
26. Supplementary data delineating the precise boundaries of the SNP blocks described in this paper as well as the haplotypes identified for each block in the 20 chromosomes sampled are available at *Science Online* at [www.sciencemag.org/cgi/content/full/294/5547/1719/DC1](http://www.sciencemag.org/cgi/content/full/294/5547/1719/DC1) and [www.perlegen.com/haplotype](http://www.perlegen.com/haplotype).

27. N. Risch, K. Merikangas, *Science* 273, 1516 (1996).
28. E. Lander, *Science* 274, 536 (1996).
29. D. Altschuler et al., *Nature Genet.* 26, 76 (2000).
30. L. Kruglyak, *Nature Genet.* 22, 139 (1999).

31. We thank B. Margus, E. Rubin, A. Chakravarti, and E. Lander for helpful discussions and an anonymous reviewer for suggestions that significantly improved the manuscript.

20 August 2001; accepted 1 October 2001

## A Genomic-Systems Biology Map for Cardiovascular Function

Monika Stoll,<sup>1</sup> Allen W. Cowley Jr.,<sup>1</sup> Peter J. Tonellato,<sup>1,2</sup> Andrew S. Greene,<sup>1</sup> Mary L. Kaldunski,<sup>1</sup> Richard J. Roman,<sup>1</sup> Pierre Dumas,<sup>1,3</sup> Nicholas J. Schork,<sup>4,5,6</sup> Zhitao Wang,<sup>1,2</sup> Howard J. Jacob<sup>1,3\*</sup>

With the draft sequence of the human genome available, there is a need to better define gene function in the context of systems biology. We studied 239 cardiovascular and renal phenotypes in 113 male rats derived from an F<sub>2</sub> intercross and mapped 81 of these traits onto the genome. Aggregates of traits were identified on chromosomes 1, 2, 7, and 18. Systems biology was assessed by examining patterns of correlations ("physiological profiles") that can be used for gene hunting, mechanism-based physiological studies, and, with comparative genomics, translating these data to the human genome.

Genetic studies of multifactorial disorders in human populations remain challenging due to the modest nature of gene effects and the heterogeneity of patient populations. The difficulties investigators face in identifying QTLs in multifactorial diseases have become apparent from the results obtained from recent total genome scans for asthma (1), hypertension (2, 3), NIDDM (4), and IDDM (5) in diverse human populations. Hypertension is one such multifactorial disorder that develops as a consequence of an "error" in the complex and redundant biological systems that determine blood pressure. The present manuscript describes the results of studies in which 239 phenotypes in each animal have been analyzed (i) to develop a model of the systems biology of the rat for renal, vascular, and neurohumoral function; (ii) to develop a correlational physiological model of the relationships among these phenotypes; and (iii) to develop bioinformatic tools to link the genetic model and the physiological model. The output for the linkage map and use of the physiological profiling tool can be found at (6).

<sup>1</sup>Department of Physiology, <sup>2</sup>Bioinformatics Research Center, and <sup>3</sup>Human and Molecular Genetics Center, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226–0509, USA. <sup>4</sup>Case Western Reserve University, Cleveland, OH 44106, USA. <sup>5</sup>The Jackson Laboratory, Bar Harbor, ME 04609, USA. <sup>6</sup>Harvard School of Public Health, Boston, MA 02115, USA.

\*To whom correspondence should be addressed. E-mail: [jacob@mcw.edu](mailto:jacob@mcw.edu)

A comprehensive genetic linkage map of 239 "likely determinant phenotypes" of blood pressure was first developed (7) using a total genome scan with a ~10-cM interval between markers to produce a detailed system biology map (6). Many of the quantitative trait loci (QTLs) for blood pressure aggregate (six or more QTLs with overlapping 95% confidence intervals) in discrete regions on rat chromosomes 1, 2, 7, and 18 (Fig. 1) (6). In four of these five aggregates, the phenotypes were independent, indicating that the cluster of traits is likely to be the result of separate genes rather than pleiotropy. In the fifth set, on chromosome 18, significant correlations were found among the phenotypes that could be divided into three functional groups, i.e., vascular reactivity, plasma lipid concentrations, and renal function, suggesting a functional genes cassette, as has been observed for QTLs in agriculture (8) and biomedical research (9–11).

To date, the majority of genome wide scans searching for the genetic basis of hypertension in rats have focused on a limited number of phenotypes, typically blood pressure and heart rate. The results of these studies have identified QTLs on almost every rat chromosome, with loci confirmed on chromosomes 1, 2, 3, 5, 10, and 13 (12). Unfortunately none of these have been translated into genes. The need for improved tools, including better phenotypes, to identify the genes responsible for these QTLs has been well articulated by Nadeau and Frankel (11). The present comprehensive linkage study in

rats addresses this problem by mapping a large number of important cardiovascular phenotypes that may serve as intermediate phenotypes or alternative traits that can be followed for positional cloning.

To assess the systems biology of these mapped cardiovascular traits, we have conducted "physiological profiling," based on ordering the phenotypes following genetic mapping. A total of 125 traits [81 parametric LOD (logarithm of the odds ratio for linkage)  $\geq 2.8$ , nonparametric LOD  $\geq 3.5$ ; 18 parametric LOD 2.5-2.8; and 26 phenotypes functionally related to blood pressure] of the 239 were mapped and incorporated into a physiological profile that visually demonstrates correlation structures using color scale for values from +1 to -1 for correlation coefficients between the traits. The genetic (first order) linkage is dependent on the direct influences of genes contributing to the mechanisms measured by the mapped traits [see Web site (6) for linkages on all chromosomes]. The physiological (second order) relationships likely reflect mechanistic relationships among pathways involved in the control of blood pressure when grouped into functionally related clusters (vascular, heart, renal, endocrine, and morphometric), and ordered within the clusters by known physiological relationships (13). Combining genetic linkage (QTL aggregates) analysis and the physiological profiles provides a way to relate genetic information with functional pathways. The present manuscript and associated Web site (6) enables the reader to access: (i) all of the results of the linkage analyses and (ii) phenotypic physiological profiles for each microsatellite marker on the linkage map sorted by genotype.

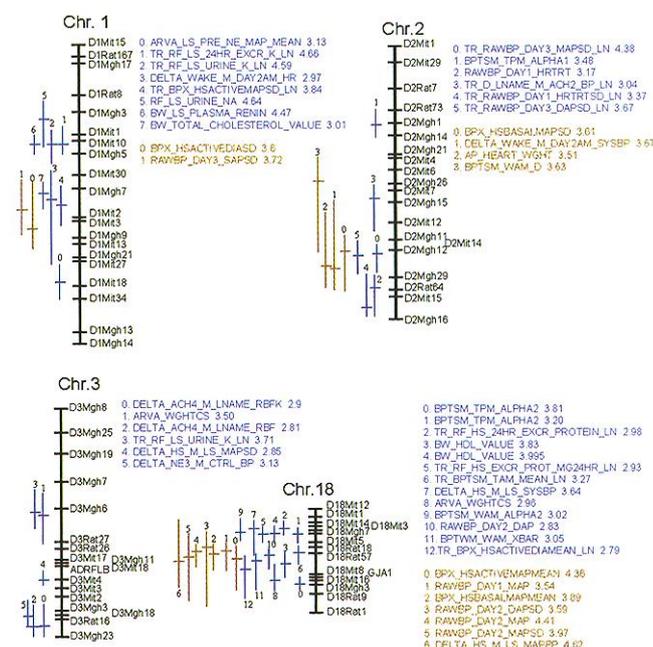
To demonstrate the utility of our approach, examples are provided for the proof of concept, as well as for relations that were not apparent via the genetic mapping data or known physiology. Linkage analysis identified suggestive QTLs for MAP and renal blood flow (RBF) responses to acetylcholine (ACh) on chromosome 10 (D10Mgh14). The QTL contributes 17% to the variance of the pressure and renal blood flow in this  $F_2$  population. In order to search for additional loci that contribute to these traits, the contribution of D10Mgh14 to the genetic variance was removed using the "fix" command in MAP-MAKER (14). Loci on chromosomes 4 (D4Mit2), and 12 (D12Mit7) contribute to the ACh response. Interestingly, all three regions harbor the genes for nitric oxide synthesis (Chr 4-NOSIII, Chr 10-NOSII, and Chr 12-NOSI). Whereas increases in the synthesis of nitric oxide mediate much of the vasodilator response to ACh (15), the response has not been associated with all three nitric oxide synthases. We therefore investigated the impact of BN and SS alleles of all three NOS

genes on the 125 profiled cardiovascular traits. Figure 2A illustrates the resulting physiological profile for the  $F_2$  male rats that were homozygous SS for D10Mgh14 (the flanking marker for NOSII), and Fig. 2B is the profile for the same marker with  $F_2$  rats homozygous for the BN allele. The correlation patterns for these two profiles are quite different. Closer scrutiny revealed positive correlations between the blood pressures related to intravenous (iv) administration of norepinephrine (NE), and angiotensin II (Ang II) (cells #85-94) in rats homozygous for the SS allele at D10Mgh14 (NOSII, Fig. 2A). Whereas  $F_2$  rats that were homozygous for the BN allele at this same marker exhibited only weak correlations among these phenotypes (Fig. 2B), the relationship of measured differences in systolic, diastolic, and mean arterial pressures before and following infusions of NE (cells #88-92) was also different. The  $F_2$  rats carrying the BN allele at NOSII (iNOS) marker D10Mgh14 exhibited a lack of correlation between arterial pressure levels before and following intravenous infusions of NE, as illustrated in Fig. 3A. Conversely, a significant positive correlation was observed in  $F_2$  rats homozygous for the SS allele at NOSII (iNOS). The arterial pressure of male rats carrying the SS allele at NOSII returned precisely to control levels, while those carrying the BN allele fell significantly below control and remained at hypotensive levels for as long as 10 min, as illustrated in Fig. 3B.

In addition, the physiological profiles for rats partitioned by alleles on chromosome 4 at D4Mit2 (NOSIII) showed similar relationships for these traits (#85-94) for SS versus BN alleles (6). Physiological profiles of two NOS isoforms from the BN contributing to the trait were also found, thus reinforcing that the presence of one or more NOS loci conferred this distinct pressure phenotype to these rats.

While the outcome of the analysis was consistent with the known importance of NO to the control of vascular tone and blood pressure, the observation of an extended hypotension in response to vasoconstrictor agents attributable to NOS is novel. A number of other relationships can be derived from this analysis that are more unexpected and serve to demonstrate how the systems biology map and physiological profiling can be used to find novel relationships and testable hypotheses. For example, in the same physiological profile presented in Fig. 2, A and B, it is seen that in  $F_2$  rats homozygous for the NOSII BN allele, there were positive correlations between the urinary excretion of protein and plasma lipid concentration (cells 115 versus 34) and between kidney weight and plasma lipid concentrations (cells 115 versus 36). Although it is well recognized that hyperlipidemia, proteinuria, and renal hypertrophy are related symptoms seen in hypertensive and diabetic nephropathy and end-stage renal disease, the influence of NOSII geno-

**Fig. 1.** Comprehensive linkage map for four of the 21 rat chromosomes. Genetic distances are calculated based on Haldane algorithm in the MAP-MAKER software package (14). These four chromosomes show clusters of likely determinant phenotypes and demonstrate that there are functional gene cassettes within the genome for these traits. Vertical bars on left side represent the 95% confidence intervals (CI) of individual QTLs. Green bars indicate CI from parametric analysis, orange bars indicate CI from nonparametric analysis. The phenotypes associated for each QTL can be found on the Web site (6), as can the remaining autosomes. Statistics for multiple comparisons: A permutation test was run 125 times to set the threshold of suggestive and significant linkage by randomly assigning genotypes with phenotype values. This permutation test found that the thresholds set by Lander and Kruglyak (19), 2.8 and 4.3, respectively, for suggestive and significant, are appropriate for this study, even with the large number of phenotypes tested. Detailed information on QTL profiles and phenotype distribution and detailed phenotype descriptions are part of the electronic supplemental material (6) of this manuscript.



REPORTS

type on the relationships between these indices of renal end organ damage and hyperlipidemia has not been described previously. Thus, the present finding that the severity of proteinuria, renal hypertrophy, and hyperlipidemia in an  $F_2$  population of rats is influenced by the allelic variations of inducible NOS (NOSII) gene is novel and creates new directions for further research.

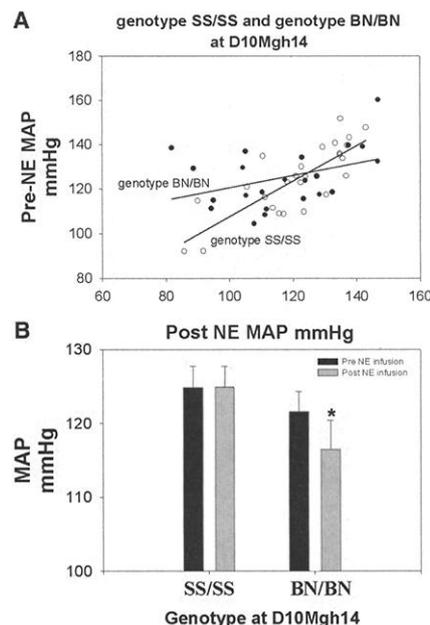
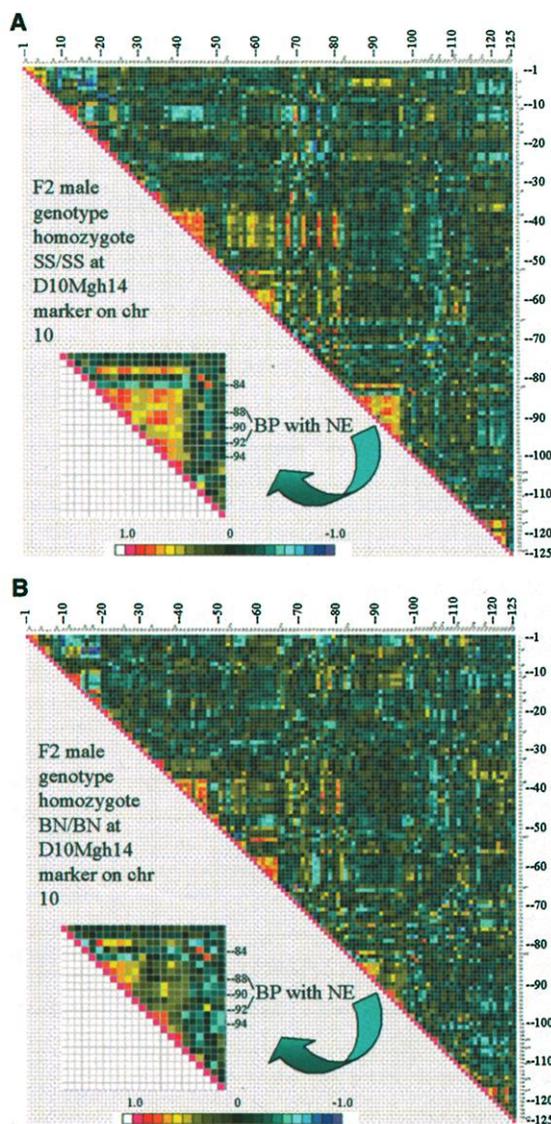
Another interesting relationship determined by physiological profiling that was not detected by linkage analyses alone was the strong positive correlation found between AngII induced reductions of renal blood flow and the chronic urinary excretion of protein in rats that were homozygous BN at the D13Mgh18 allele. Conversely, a significant negative correlation was found in  $F_2$  rats that were homozygous SS at this same allele. Urine protein excretion was determined during steady-state conditions of high salt intake. Renal blood flow responses to three doses of AngII were determined following several

days of salt depletion in anesthetized  $F_2$  rats. These results are of considerable interest because they provide a genetic basis for the results of many clinical studies of essential hypertension over the past 30 years, in which patients have been stratified based on their renal vascular responses to AngII. In 40 to 50% of the essential hypertensive population, adrenal and renal vascular responses to AngII are not modified as expected by changes in sodium intake and these individuals have been called "non-modulators" (16). It has been documented that nonmodulators exhibit a higher percentage of one or both parents with hypertension, suggesting that this renal abnormality is inherited and linked to the development of hypertension (16). In the present study, SS fed a high salt diet exhibited substantial proteinuria compared to BN parental rats. The physiological profile revealed that in our  $F_2$  rats, it was possible to predict the renal blood flow AngII sensitivity based on genotype and protein excretion lev-

els. A narrow region on rat chromosome 13 near D13Mgh18 enables a prediction of modulators and nonmodulators using individual genotype. The only known gene with a defined function consistent with the AII response that maps very closely to D13Mgh18 is renin. However, this does not mean that renin is the causative gene, but rather provides a hypothesis that can be tested.

In summary, we have developed a comprehensive systems biology map for cardiovascular traits in an  $F_2$  population of rats and a physiologic profiling tool to assess the complex relationships between these phenotypes as a function of genotype. Several examples have been provided that demonstrate the utility of this analytical approach to generate new hypotheses regarding gene and functional relationships. Physiological profiling can reveal functional interactions between traits and the genome that are not apparent in QTL linkage analysis of independent traits. The finding that mechanistically related cardiovascular pheno-

**Fig. 2.** Physiological profiles (A and B) each represent a correlation matrix with phenotypic ordering of 125 likely determinants of arterial blood pressure, based on Guyton's model of blood pressure control, on each axis (73). The utilization of these profiles enables one to assess the impact of allelic substitutions on each of these traits in either the parental, or  $F_2$  generation rats of the SS/JrHsd/Mcw  $\times$  BN/SsNHsd/Mcw intercross. The scale shows that strong positive correlations are represented in red, strong negative correlations are blue, while low correlations are in gray and black. Profile (A) represents  $F_2$  male rats that were homozygous SS for D10Mgh14 (NOSII gene region), and profile (B) for the same marker with  $F_2$  rats homozygous for the BN allele. As shown by the expanded insert, there was a high degree of correlation between blood pressures determined immediately before, during, and after the short-term infusions of norepinephrine (NE), angiotensin II, and acetylcholine in  $F_2$  rats homozygous SS for the NOSII gene. In contrast, profile (B) shows that these correlations are virtually lost in  $F_2$  rats that were homozygous BN for the NOSII locus. Investigators interested in physiological genomics can electronically access our Web site (6) to examine profiles related to their own particular cardiovascular interests for any of the microsatellite markers shown in the genetic map of Fig. 1.



**Fig. 3.** Illustration of the relationship between mean arterial pressure (MAP) before and after infusions of norepinephrine (NE) in  $F_2$  rats homozygous SS for the NOSII gene and those homozygous BN for the NOSII gene. The graph in (A) demonstrates that the correlation between the pre- and postinfusion of NE was greater in  $F_2$  rats that were homozygous SS (open circles) at the marker for the NOSII gene than in  $F_2$  rats that were homozygous BN (closed circles) for NOSII. (B) shows the average levels of MAP before (solid bars) and following completion (open bars) of the iv infusions of three doses of norepinephrine (NE). It is seen that MAP of male rats carrying the SS allele at NOSII returned precisely to control levels, while those carrying the BN allele fell significantly below control. The heterozygote at D10Mgh14 (18) is intermediate between SS and BN, but lies closer to BN.

types aggregate in QTL regions, combined with physiological profiling, provides a novel approach to facilitate the positional cloning of genes underlying cardiovascular function and hypertension. While strategies for accomplishing these goals will continue to evolve, this first attempt at developing knowledge at a systems biology level sets the stage for future improvements and provides investigators with a powerful set of tools for discovery.

References and Notes

1. E. R. Bleecker, D. S. Postma, D. A. Meyers, *Am. J. Respir. Crit. Care Med.* **156**, S113 (1997).
2. C. Julier *et al.*, *Hum. Mol. Genet.* **6**, 2077 (1997).
3. J. Krushkal *et al.*, *Circulation* **99**, 1407 (1999).
4. C. L. Hanis *et al.*, *Nature Genet.* **13**, 161 (1996).
5. J. A. Todd, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8560 (1995).
6. A user interface to the complete data set is found at <http://brc.mcw.edu/phyprf>.
7. The total genome scan was carried out with an average 10 cM spacing of markers following phenotyping with 239 measured or derived traits. After testing for normalcy, 166 traits were analyzed in a parametric genome scan using MAPMAKER/QTL (12, 16). The remaining 73 traits were analyzed using the nonparametric mapping algorithm (MAPMAKER/QTL version 1.9b). (Phenotyping protocol, conscious): All blood pressure measurements were made with the animals unrestrained in their home cage as described previously (17). Data were collected at a rate of 100 Hz and reduced to 1-min averages, except for time series analysis where they were 1-s averages. The day-night light cycle for all rats ran from 2:00 a.m. (lights on) to 2:00 p.m. (lights off) throughout the study. BP1 (High Salt Day 1): Baseline measurements of systolic, diastolic, and mean arterial pressure and heart rate were measured from 9:00 a.m. to noon. BP2 (High Salt Day 2 – inactive versus active): Repeat morning recording ("inactive phase," lights on), and record 4 hours during the dark cycle (2:00 p.m. to 6:00 p.m. "active phase"). Urine was collected for 24 hours for measurement of volume, sodium, potassium, protein, and creatinine. All BP data this day was collected for time-series analysis. BP3 (High Salt Day 3): Repeat morning recording. Following the recording period, a blood sample (500 µl) was drawn for determination of creatinine, plasma renin activity, plasma protein, and hematocrit. Furosemide challenge: Following blood draw, an ip injection of furosemide (10 mg/kg) was given to salt deplete the animals, and the diet switched to a 0.4% low salt. BP4 (Stress test): Following a control period, an alerting stimulus (2 mA for 0.3 s) was delivered twice with a 5-min interval during pressure recording. Change in mean arterial pressure, the time to peak, and the time to 90% recovery were determined. BP5 (Low Salt): Repeat 3-hour morning recording of blood pressure in the salt depleted state. Following the recording period, a 1.0-ml blood sample was taken for the measurement of plasma renin activity, triglycerides, total cholesterol, HDL, creatinine, hematocrit, and white blood cell count. (Phenotyping protocol, anesthetized): Rats were anesthetized with ketamine (30 mg/kg; intramuscular) and Inactin (50 mg/kg; intraperitoneal). Catheters were implanted in the femoral artery and vein and an electromagnetic flow probe was placed on the left renal artery via a midline incision. An iv infusion (50 µl/min) of isotonic saline containing 1% bovine serum albumin replaced fluid losses. After a 45-min equilibration, arterial blood pressure and renal blood flow were measured during a 15-min control period. Renal and peripheral vascular responses to 5-min iv infusions of angiotensin II (20, 100, 200 ng/kg/min) and norepinephrine (0.5, 1, 3 µg/kg/min) were determined. Following recovery of pressure to baseline values, renal vascular and systemic arterial responses to two successive doses of acetylcholine (ACh) (0.1 and 0.2 µg/kg/min) were measured. L-NAME (5 mg/kg) was then administered

as an iv bolus to determine the contribution of nitric oxide to basal renal vascular tone. After 10 min of equilibration, a repeat infusion of the same two doses of ACh were administered to test for the degree of blockade of the synthesis of nitric oxide produced by L-NAME. (Morphometric measurements): Heart and kidneys were removed, stripped of surrounding tissue, and weighed to assess the degree of cardiac and renal hypertrophy. (Histology): The right kidney was immersion-fixed in 10% buffered formalin and embedded in paraffin, and prepared sections stained with H&E and PAS were evaluated for mean glomerular diameter and the degree of focal glomerulosclerosis.

8. B. R.Thumma *et al.*, *J. Exp. Bot.* **52**, 203 (2001).
9. L. L. Miner *et al.*, *Psychopharmacology* **117**, 62 (1995).
10. K. G. Becker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9979 (1998).
11. J. Nadeau, W. Frankel, *Nature Genet.* **25**, 381 (2000).
12. J. P. Rapp, *Physiol. Rev.* **80**, 135 (2000).
13. A. C. Guyton *et al.*, *Annu. Rev. Physiol.* **34**, 13 (1972).

14. E. S. Lander *et al.*, *Genomics* **1**, 174 (1987).
15. J. Loscalzo, G. Welch, *Progr. Cardiovasc. Dis.* **38**, 87 (1995).
16. N. K. Hollenberg *et al.*, *Medicine* **57**, 167 (1978).
17. A. W. Cowley Jr. *et al.*, *Physiol. Genomics* **2**, 107 (2000).
18. M. Stoll, A. W. Cowley Jr., A. S. Greene, data not shown.
19. E. S. Lander, L. Kruglyak, *Nature Genet.* **11**, 241 (1995).
20. We thank the Bioinformatics Research Center at Medical College of Wisconsin for the development of the database, tools, and Web site; M. Granados, M. Nobrega, M. Shiozawa, and M. Runte for assistance with genotyping; A. Kwitek-Black for comparative mapping; and T. Kurth, K. Bork, P. Regozzi, and C. Thomas for excellent technical assistance. This work was supported by National Heart, Lung, and Blood Institute grant 1P50-HL-54998.

1 May 2001; accepted 17 October 2001

## Estimation of Epidemic Size and Incubation Time Based on Age Characteristics of vCJD in the United Kingdom

Alain-Jacques Valleron,<sup>1</sup> Pierre-Yves Boelle,<sup>1</sup> Robert Will,<sup>2</sup> Jean-Yves Cesbron<sup>3</sup>

The size of the variant Creutzfeldt-Jakob Disease (vCJD) epidemic in the United Kingdom is a major public health concern and a subject of speculation. The cases are young (mean age = 28). Assuming that the risk of developing the disease in susceptible exposed subjects decreases exponentially with age after age 15, that all infections occurred between 1980 and 1989, and that the distribution of the incubation period is lognormal, we estimate that the mean duration of the incubation period is 16.7 years [95% confidence interval (CI): 12.4 to 23.2] and that the total number of cases will be 205 (upper limit of the 95% CI: 403).

As of 1 May 2001, there were 97 cases of definite ( $n = 86$ ) or probable ( $n = 11$ ) variant Creutzfeldt-Jakob Disease (vCJD) in the United Kingdom. These patients probably contracted the disease by oral ingestion of food contaminated by the agent of bovine spongiform encephalopathy (BSE), before the UK bovine-specified risk materials (SRM) ban in 1989 (1). The number of BSE-infected animals is estimated to have been in the range of 900,000 to 1,130,000, with between 460,000 and 482,000 slaughtered for consumption before the introduction of the November 1989 specified offal ban (2). The epidemic may have started as early as 1980 (3), and the number of people exposed to potentially infective doses through food may

be extremely high (4). Therefore, one could pessimistically assume that virtually everybody in the population has been in contact with food, or bovine products, originating from BSE-infected animals. The public health response in Europe has been to develop procedures and diagnostic tests that avoid, as far as possible, the entry of any infected animal into the food chain. The consequences of the BSE epidemic in terms of human disease are not yet known: With different assumptions for risk analysis, in 1997, the cumulative cases of vCJD in the United Kingdom were estimated from as few as 75 to as many as 80,000 (5) and more recently from 70 to 136,000 cases (6). These estimates are markedly dependent on assumptions made about the mean duration of the incubation period. Unfortunately, no studies in animals or of other human spongiform encephalopathies provide precise data for the incubation period. In addition, the observation that only a few cattle, and often only one, from the same age cohort in a herd have developed BSE suggests that additional individual or environmental factors may influence the development of the disease. It is be-

<sup>1</sup>Epidemiology and Information Sciences, INSERM U444, CHU Saint-Antoine, Université Pierre et Marie Curie et Assistance Publique-Hôpitaux de Paris, 27 rue Chaligny, 75012 Paris, France. <sup>2</sup>National Creutzfeldt-Jakob Disease Surveillance Unit, Western General Hospital, Edinburgh EH4 2XU, UK. <sup>3</sup>Immunité Anti-Infectieuse JE 2236, UFR de Médecine de Grenoble, Université Joseph Fourier, Domaine de la Merci, 38706 La Tronche, France.