HA on endothelium induced by proinflammatory stimuli (*30*), this ligand pair would participate in the process of extravasation into these sites (*8*).

REFERENCES AND NOTES

1. C. R. Mackay, *Immunol. Today* **12**, 189 (1991).
2. F. T. Koster and D. D. McGregor, *J. Exp. Med.* **133**, 864 (1971).
3. G. D. Guy, C. Griscelli, P. Vassalli, *ibid.* **148**, 1661 (1978).
4. T. B. Issekutz, *J. Immunol.* **147**, 4178 (1991).
5. E. C. Butcher and L. J. Picker, *Science* **272**, 60 (1996).
6. T. A. Springer, *Cell* **76**, 301 (1994).
7. J. Lesley, R. Hyman, P. W. Kincade, *Adv. Immunol.* **54**, 271 (1993).
8. H. C. DeGrendele, P. Estess, L. J. Picker, M. H. Siegelman, *J. Exp. Med.* **183**, 1119 (1996).
9. J. Lesley, N. Howes, A. Perschl, R. Hyman, *ibid.* **180**, 383 (1994).
10. H. C. DeGrendele, P. Estess, M. H. Siegelman, *J. Immunol.* **159**, 2549 (1997).
11. M. T. Scherer, L. Ignatowicz, G. M. Winslow, J. W. Kappler, P. Marrack, *Annu. Rev. Cell Biol.* **9**, 101 (1993).
12. T. Herrmann, S. Baschieri, R. K. Lees, H. R. MacDonald, *Eur. J. Immunol.* **22**, 1935 (1992).
13. J. E. McCormack, J. E. Callahan, J. Kappler, P. C. Marrack, *J. Immunol.***150**, 3785 (1993).
14. J. Sprent, J. F. A. P. Miller, G. F. Mitchell, *Cell. Immunol.* **2**, 171 (1971).
15. R. N. Cahill, H. Frost, Z. Trnka, *J. Exp. Med.* **143**, 870 (1976).
16. J. Sprent, *Cell* **76**, 315 (1994).
17. K. Miyake, C. B. Underhill, J. Lesley, P. W. Kincade, *J. Exp. Med.* **172**, 69 (1990).
18. M. H. Julius, E. Simpson, L. A. Herzenberg, *Eur. J. Immunol.* **3**, 645 (1973). Enrichment of T lymphocytes from blood, lymph node, and peritoneal samples was done by nylon wool nonadherence. Resulting cell samples contained >80%, >95%, and >80% T cells, respectively, based on flow cytometric analysis with antibody to $TCR_{\alpha\beta}$. T cell depletion of peripheral blood and peritoneal samples was done with anti-Thy1.2 conjugated magnetic beads (Dynal). Cells (5×10^6) were incubated for 20 min with 10^7 beads in 1 ml of RPMI 1640 with 5% fetal bovine serum (FBS) at 4°C on a rocking platform. $Thy1.2^+$ T cells were removed by magnetic separation. After depletion, <1% of the remaining cells were shown to be T cells by flow cytometric analysis. Depletions of $V_\beta8^+$ T cells and HA-binding cells were done with anti-$V_\beta8$–biotin or Fl-HA + anti-fluorescein isothiocyanate (FITC)–biotin, respectively, and then with streptavidin-conjugated magnetic beads (Dynal). Depletion of $V_\beta14^+$ T cells was done with anti-$V_\beta14$ and GαRIg-conjugated beads (Dynal). Control samples were incubated with streptavidin-conjugated beads alone. For positive selection of $V_\beta8^+$ T cells, cells were incubated with anti-$V_\beta8$–biotin in the presence of 10% normal mouse serum, washed, and then incubated with streptavidin-FITC. After washing twice, cells were further incubated with anti-FITC magnetic particles and separated from the total population using a MACS magnetic separation column (Miltenyi Biotec). Fluorescence-activated cell sorting (FACS) analysis of this population showed them to be 85 to 90% $V_\beta8^+$. Positive selection of $V_\beta14$ T cells was done with anti-$V_\beta14$, anti-κ–FITC, and anti-FITC particles.
19. H. C. DeGrendele, P. Estess, M. H. Siegelman, data not shown.
20. For short-term homing assays, cells were labeled by resuspending at a concentration of 10^7 cells/ml in Hanks' balanced salt solution (HBSS) containing 2 μM 5- (and 6-)carboxylfluorescein diacetate, succinimidyl ester (CFDA, Molecular Probes), incubating at room temperature for 20 min, then washing twice in HBSS. Recipient BALB/c mice [injected ip 20 hours earlier with SEB (50 μg) or with PBS] were then injected iv with 10^7 labeled cells per mouse in 0.5 ml of HBSS with or without the addition of blocking reagent. Depletions were done as above. Cells treated with whole antibody or Fab fragments were incubated in a saturating concentration of antibody for 15 min on ice and then washed before injection. HA and chondroitin sulfate (Sigma) were added to the cell suspension at the time of injection (final concentration, 0.5 μM). Hyaluronidase (ICN Biochemicals) and chondroitinase ABC (Sigma) treatment was done by injecting 10 U per mouse iv 30 min before donor cell infusion. Ninety minutes after the infusion of labeled cells, the recipient mice were killed and PELs were analyzed for the homing of labeled cells.
21. G. F. Weber, S. Ashkar, M. J. Glimcher, H. Cantor, *Science* **271**, 509 (1996).
22. P. R. Maroko *et al.*, *N. Engl. J. Med.* **296**, 898 (1977).
23. L. Szlavy, S. Kubik, A. deCourten, H. J. Hachen, F. Solti, *Angiology* **36**, 452 (1985).
24. S. T. Jalkanen, R. F. Bargatze, L. R. Herron, E. C. Butcher, *Eur. J. Immunol.* **16**, 1195 (1986).
25. S. Jalkanen, R. F. Bargatze, J. de los Toyos, E. C. Butcher, *J. Cell Biol.* **105**, 983 (1987).
26. R. L. Camp, A. Scheynius, C. Johansson, E. Pure, *J. Exp. Med.* **178**, 497 (1993).
27. K. Mikecz, F. R. Brennan, J. H. Kim, T. T. Glant, *Nature Med.* **1**, 558 (1995).
28. B. F. Haynes, L. P. Hale, K. L. Patton, M. E. Martin, R. M. McCallum, *Arthritis Rheum.* **34**, 1434 (1991).
29. A. Zeidler *et al.*, *Autoimmunity* **21**, 245 (1995).
30. M. Mohamadzadeh, H. C. DeGrendele, P. Estess, M. H. Siegelman, in preparation.
31. BALB/c mice were injected ip with 50 μg of SEB (Sigma) in 500 μl of sterile PBS and killed at 4-hour intervals for analysis. Controls were injected with PBS alone. MLNs were processed into single-cell suspensions and the percentage of $V_\beta8^+$ lymphocytes was determined by flow cytometry using anti-$V_\beta8$-phycoerythrin (PE) (Pharmingen). Peripheral blood was collected by cardiac puncture into Alsever's solution and red blood cells were lysed by incubation in 0.15 M NH_4Cl, 1 mM $KHCO_3$, and 0.1 mM EDTA. PELs were collected by peritoneal lavage using 5 ml of RPMI (37°C) containing 2% FBS and 2 mM EDTA. Staining of PELs was performed as above, except cells were preincubated with anti-CD32/CD16 (clone 2.4G2) and stained in the presence of 10% normal mouse serum to inhibit Fc receptor interactions. Data were collected on a FACScan analytical instrument (Becton Dickinson) and analyzed using Lysis II software.
32. A. N. de Belder and K. O. Wik, *Carbohydr. Res.* **44**, 251 (1975).
33. D. A. Jones, O. Abbassi, L. V. McIntire, R. P. McEver, C. W. Smith, *Biophys. J.* **65**, 1560 (1993). For blocking studies, antibodies were added at saturating concentrations to the cell suspension before their addition to the flow system. Rolling cells were scored visually and are reported as the cumulative number of interacting cells by scanning the entire surface of the chamber (0.9 cm^2) from outlet to inlet over a 4-min period.
34. R. Kennett, *Monoclonal Antibodies* (Plenum, New York, 1980).

35. We thank M. Kosfiszer for technical assistance and L. Picker for helpful discussions. Supported by grants from NIH (R01 CA57471 and HL56746), the Arthritis Foundation, and the Welch Foundation (I-1227). M.H.S. is an Established Investigator of the American Heart Association. H.C.D. is supported by an NIH Cancer Immunology Training Grant.

8 July 1997; accepted 18 September 1997

Functional Coherence of the Human Y Chromosome

Bruce T. Lahn and David C. Page

A systematic search of the nonrecombining region of the human Y chromosome (NRY) identified 12 novel genes or families, 10 with full-length complementary DNA sequences. All 12 genes, and six of eight NRY genes or families previously isolated by less systematic means, fell into two classes. Genes in the first group were expressed in many organs; these housekeeping genes have X homologs that escape X inactivation. The second group, consisting of Y-chromosomal gene families expressed specifically in testes, may account for infertility among men with Y deletions. The coherence of the NRY's gene content contrasts with the apparently haphazard content of most eukaryotic chromosomes.

Functional or developmental themes have rarely been ascribed to whole chromosomes in eukaryotes. Instead, individual chromosomes appear to contain motley assortments of genes with extremely heterogeneous patterns of developmentally regulated expression. We speculated that the human Y chromosome might be a functionally coherent exception, at least in its nonrecombining portion (the NRY), which makes up 95% of its length (*1*). It is known to differ from all other nuclear human chromosomes by the absence of recombination, its presence in males only, its common ancestry and persistent meiotic relationship with the X chromosome, and the tendency of its genes to degenerate during evolution (*2*).

From the 1950s to the present day, many biologists have assumed that the Y chromosome is a functional wasteland, despite the discovery of several NRY genes during this period. Studies of human pedigrees had identified many traits exhibiting autosomal or X-linked inheritance but no convincing cases of Y-linked inheritance (*3*). In 1959, reports of XO females and XXY males established the existence of a sex-determining gene on the human Y chromosome (*4*), but this was perceived as a special case on a generally desolate chromosome. The wasteland model has been revised only during the past decade, when eight NRY transcription

Howard Hughes Medical Institute, Whitehead Institute, and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA.

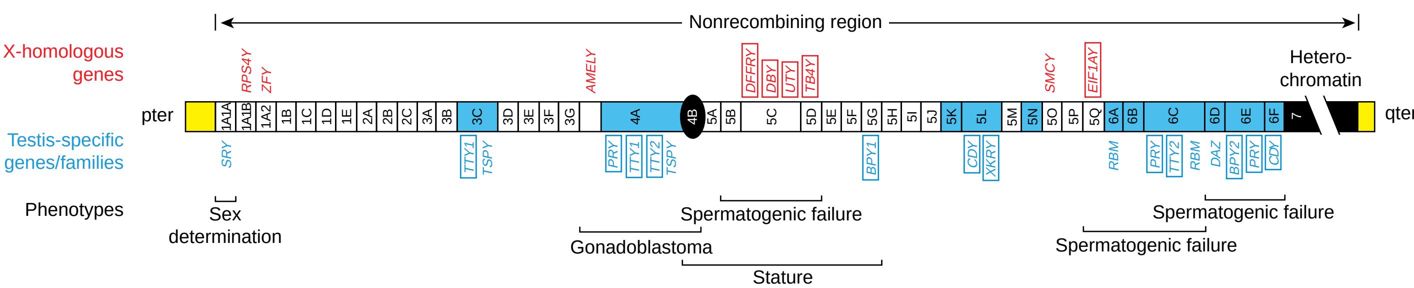

Fig. 1. Gene map of NRY. The Y chromosome consists of a large nonrecombining region (NRY; euchromatin plus heterochromatin) flanked by pseudoautosomal regions (yellow). Pter, short-arm telomere; qter, long-arm telomere. The NRY is shown divided into 43 ordered intervals (1A1A through 7) defined by naturally occurring deletions; deletion intervals previously shown to contain Y-specific repeats are shaded blue (*10*, *11*). Listed immediately above the chromosome are nine NRY genes with functional X homologs (red); novel genes are boxed. Immediately below the chromosome are 11 testis-specific genes or families (blue), some with multiple locations. Within deletion intervals, genes have not been ordered. Some testis-specific families probably have members in additional deletion intervals; indicated locations are representative but not necessarily exhaustive. At bottom are shown NRY regions implicated, by deletion mapping, in sex determination, germ cell tumorigenesis (gonadoblastoma), determination of stature, and spermatogenic failure (*7*, *8*, *28*, *31*). For *DFFRY*, previously thought to be a pseudogene, these mapping studies confirm published findings (*19*).

units (or families of closely related transcription units) were identified, mostly during regionally focused, positional cloning experiments (*5*–*8*). Even in recent years, it has been argued that the NRY's gene content is essentially limited to random disintegrating vestiges of its common ancestry with the X (*9*). The Y-specific repetitive sequences that are so plentiful in the euchromatic regions (*10*, *11*) have often been assumed to be functionally inert (*12*). Realizing that these wasteland theories were based on limited anecdotal data about the NRY's gene content, we decided to embark on a broad, systematic gene hunt that could uncover previously unrecognized functional patterns.

A complete description of the NRY's gene content cannot be obtained with current research methods, short of sequencing the entire NRY. However, it should be possible to obtain a broad representative sampling of NRY genes that could enable us to make comprehensive generalizations. We searched for this sampling in sequences transcribed in a single complex organ, the testis. To assess the suitability of the testis and of a "cDNA selection" protocol (*13*) for this project, we first sought to crudely measure what fraction of human genes, regardless of developmental regulation, are detectably transcribed there. We did this by testing whether previously identified pseudoautosomal genes (*1*), whose diversity in developmentally regulated expression is like that of autosomal genes, could be found among testis transcripts. The nine known pseudoautosomal genes were previously identified using mRNA sources as specialized as liver, pineal gland, and skeletal muscle. The extent to which we recovered the nine known pseudoautosomal genes from sampling of testis cDNA would provide a measure of this tissue's adequacy in representing a broad array of genes.

In fact, we recovered testis cDNAs for all nine known pseudoautosomal genes, which suggested that the testis as a single source would be sufficient to provide nearly comprehensive access to NRY genes. From primary, uncloned testis cDNA, we selected and determined the nucleotide sequence of 2539 fragments that hybridized to Y-chromosomal DNA. We anticipated that these sequence fragments would represent a redundant sampling of a much smaller set of genes. Nucleotide sequence analysis revealed that 579 fragments corresponded to known Y genes, including all nine pseudoautosomal genes previously reported and seven of eight known NRY genes. (The one previously reported NRY gene that we failed to recover was *AMELY*, which is expressed only in developing tooth buds.) After further analysis, both to eliminate human repetitive sequences and to assemble overlapping fragments into contigs, novel sequences were hybridized to Southern (DNA) blots of human genomic DNAs. Sequences that detected at least one prominent male-specific fragment were judged likely to derive from the NRY, and for each we attempted to isolate cDNA clones from a human

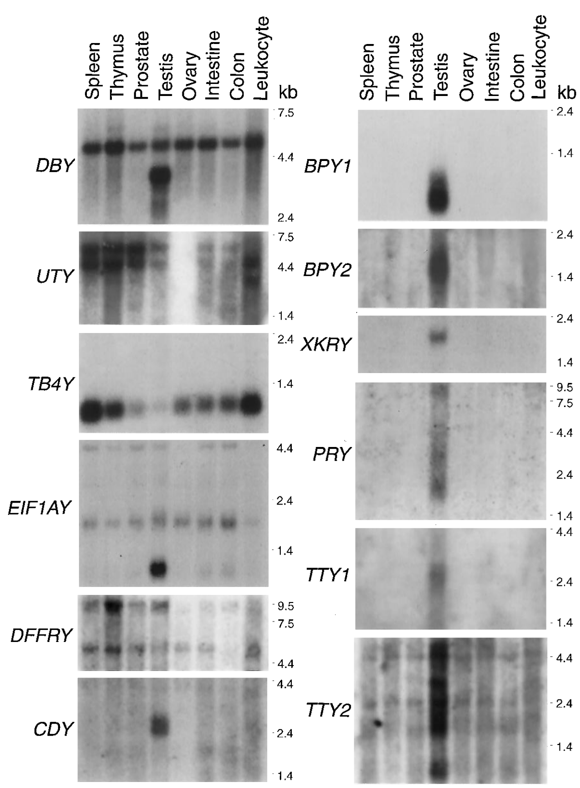

Fig. 2. Transcription of 12 novel NRY genes in human tissues. Autoradiograms were produced by hybridization of ^{32}P-labeled cDNA probes to Northern blots of polyadenylated RNAs (2 μg per lane) from human tissues (Clontech). Probes used were cDNA clones, either full-length (most genes) or partial (*DBY*, nucleotides 1476 through 2319 of GenBank no. AF000985; *UTY*, nucleotides 861 through 1768 of GenBank no. AF000996; and *DFFRY*, nucleotides 8604 through 9878 of GenBank no. AF000986). Hybridization was done at 65°C in 0.5 M Na_iPO_4 (pH 7.5) and 7% SDS; washing was done at 65°C in 1× SSC and 0.1% SDS. *DBY*, *TB4Y*, *EIF1AY*, and *DFFRY* probes cross-hybridize to transcripts derived from their X homologs. For all five X-homologous genes (*DBY*, *PRY*, *TB4Y*, *EIF1AY*, and *DFFRY*), expression was tested and confirmed in three male tissues (brain, prostate, and testis) by RT-PCR using Y-specific primers (not shown in figure). For *DFFRY*, previously thought to be a transcribed pseudogene, these expression studies confirm published findings (*19*).

Table 1. Twelve novel genes and gene families in the NRY.

NRY genes and gene families						Functional X homologs			
Gene symbol	Gene name	Comments	GenBank no.	Tissue expression	Multicopy on Y	Gene symbol	GenBank no. or ref.	X-Y amino acid sequence identity	Escape X inactivation
DBY	*Dead Box Y*	Novel protein; "DEAD box" motif suggests that this may be an RNA helicase (*32*)	AF000985 AF000984	Ubiquitous	No	*DBX*	AF000983 AF000982	91%	Yes
TB4Y	*Thymosin β4, Y isoform*	X homolog sequesters actin (*17*)	AF000989	Ubiquitous	No	*TB4X*	(*17*)	93%	Yes
EIF1AY	*Translation Initiation Factor 1A, Y isoform*	X homolog is an essential initiation factor (*18*)	AF000987	Ubiquitous	No	*EIF1AX*	(*18*)	97%	Yes
UTY	*Ubiquitous TPR motif Y*	Mouse Y homolog recently shown to encode an H-Y antigen; contains 10 tandem "TPR" motifs implicated in protein-protein interaction (*33*); differential splicing may generate isoforms differing at COOH-terminals	AF000996 AF000995 AF000994	Ubiquitous	No	*UTX*	AF000992 AF000993	85%	Yes
DFFRY	*Drosophila Fat Facets Related Y*	X homolog recently described; Y previously thought to carry a transcribed pseudogene; homologous to *Drosophila* deubiquinating enzyme required for eye development and oogenesis (*19, 34*)	AF000986	Ubiquitous	No	*DFFRX*	(*19*)	91%	Yes
CDY	*Chromodomain Y*	Novel protein with "chromodomain" (*35*) and putative catalytic domain (*36*); might modify DNA or chromosomal proteins during spermatogenesis	AF000981	Testis	Yes				
BPY1	*Basic Protein Y 1*	Novel 125-residue protein rich in Ser, Lys, Arg, and Pro calculated; isoelectric point (pI) 9.4; Southern blotting reveals X homolog, but no X-derived cDNA clones identified to date	AF000979	Testis	Yes				
BPY2	*Basic Protein Y 2*	Novel 106-residue protein; calculated pI 10.0	AF000980	Testis	Yes				
XKRY	*XK Related Y*	Novel protein with similarity to XK, a putative membrane transport protein (*37*)	AF000997	Testis	Yes				
PRY	*PTP-BL Related Y*	Novel protein with some similarity to PTP-BL, a putative protein tyrosine phosphatase (*38*)	AF000988	Testis	Yes				
TTY1	*Testis Transcript Y 1*	No significant open reading frame identified	AF000990	Testis	Yes				
TTY2	*Testis Transcript Y 2*	No significant open reading frame identified	AF000991	Testis	Yes				

testis library (*13*). Nucleotide sequencing of cDNA clones and rescreening of libraries as necessary yielded full-length cDNA sequences for 10 novel NRY genes or families and partial cDNA sequences for two additional ones (Table 1). We localized all 12 novel genes on the Y chromosome (Fig. 1) (*14*) and assessed their expression in diverse human tissues by Northern (RNA) blotting (Fig. 2). The novel genes encode an assortment of proteins (Table 1) and are dispersed throughout the euchromatic portions of the NRY (Fig. 1).

Although our gene hunt was systematic, it is likely that some NRY genes in addition to *AMELY* escaped detection; this could have resulted from failure to select corresponding cDNAs or from discarding them during subsequent screening steps. Like *AMELY*, other NRY genes may not have been recovered because they are not transcribed in sufficient amounts in the testis. Our screening criteria may have discriminated against NRY genes located in regions of exceptionally high sequence similarity to the X chromosome. In particular, we may have overlooked genes located in a 4-Mb region of the NRY characterized by 99% sequence identity to the X (*15*). Nonetheless, we suspect that the gene hunt was sufficiently comprehensive for us to form meaningful generalizations about the NRY's gene content.

The 12 novel genes readily sort into two discrete classes (Table 1). The first group, of five novel NRY genes, has several shared features. Each gene has a homolog on the X chromosome encoding a very similar but nonidentical protein isoform; every gene is expressed in a wide range of human tissues; and each gene appears to exist in a single copy on the NRY. The other seven novel NRY genes constitute the second group and share quite different traits. They appear to be expressed specifically in testes. They also seem to exist in multiple copies on the NRY, as judged by (i) the number and intensity of hybridizing fragments on genomic Southern blots or (ii) multiple map locations on the Y. The two classes of genes suggested by our NRY-wide search also accommodate six of eight NRY genes previously identified by less systematic means (*5–8*, *16*), confirming the validity of this bipartite classification.

Many of the X-homologous genes appear to be involved in cellular housekeeping, as suggested by their ubiquitous expression and by the functions of their encoded proteins, which are well established in three cases. *TB4Y* encodes a Y isoform of thymosin β_4, which functions in actin sequestration (*17*) and which we found to be encoded by the X chromosome. *EIF1AY* encodes a Y isoform of eIF-1A, an essential translation initiation factor (*18*). *RPS4Y* encodes a Y isoform of an essential ribosomal protein (*6*).

In contrast with these single-copy, X-homologous housekeeping genes, the multicopy NRY gene families appear to encode proteins with more specialized functions. All appear to be expressed specifically in the testis. Our study identified full-length cDNA clones for five of these gene families, which were all found to encode proteins not previously characterized (Table 1). Several of the testis-specific gene families may encode DNA- or RNA-binding proteins, including two small, unrelated basic proteins, BPY1 and BPY2 (Table 1); two putative RNA-binding proteins, RBM and DAZ (*7*, *8*); and CDY, which contains a "chromodomain" [a chromatin-binding motif (Table 1)].

We postulate that the NRY's evolution was dominated by two strategies. The first strategy favored conservation of particular X-Y gene pairs to maintain comparable expression of certain housekeeping functions in males and females. This strategy is at odds with the general behavior of X-Y gene pairs during mammalian evolution. The mammalian X and Y chromosomes evolved from autosomes; most ancestral gene functions were retained on the nascent X chromosome but deteriorated on the nonrecombining portion of the emerging Y chromosome (*2*). This resulted in females having two copies but males having only one copy of many genes, an inequality predominantly addressed in mammals by transcriptional silencing, or inactivation, of one X chromosome in females. Our findings on X-homologous NRY genes, together with previous studies, suggest the importance in human evolution of an additional solution: preservation of homologous genes on both NRY and X, with male and female cells expressing two copies of such genes. A critical prediction of this model is that the X homologs should escape X inactivation. This is the case for all widely expressed X-linked genes with known NRY homologs, including the X homologs of the five novel NRY genes reported here (*6*, *19–21*). A second prediction is that the X- and Y-encoded proteins should be functionally interchangeable despite considerable divergence of their genes' nucleotide sequences. Indeed, each of the eight known X-NRY gene pairs encodes closely related isoforms, with 85 to 97% amino acid identity throughout their lengths; functional interchangeability has been demonstrated in the one case tested to date (*22*).

These dosage compensation strategies may be relevant to Turner syndrome, which is classically associated with an XO sex chromosome constitution. The Turner phenotype may be due to inadequate expression of certain X-Y common genes that escape X inactivation (*23*). Given that several X-NRY genes appear to be involved in cellular housekeeping, we speculate that some features of the XO phenotype (such as poor fetal viability) reflect inadequate expression of particular housekeeping functions. The X-homologous NRY genes (Fig. 1) should be investigated as Turner candidates (*24*).

In addition to the strategy for conserving certain X-Y gene pairs, a second strategy probably shaped the NRY's evolution. This strategy favored the acquisition of testis-specific families, perhaps through selectively retaining and amplifying genes that enhance male reproductive fitness. Animal genomes may contain genes or alleles that enhance male reproductive fitness but are inconsequential or even detrimental to females, as first appreciated by Fisher (*25*). Fisher recognized that selective pressures would favor the accumulation of such genes in male-specific regions of genomes. Of course, male reproductive fitness depends critically on sperm production, the task of the adult testis. As the only male-specific portion of the mammalian genome, the NRY should have a unique tendency to accumulate male-benefit genes during evolution. Consider the human NRY's *DAZ* gene cluster, de novo deletions of which are associated with severe spermatogenic defects (*8*). The *DAZ* cluster on the human Y chromosome arose during primate evolution by transposition and amplification of an autosomal gene. Similarly, two other testis-specific NRY gene families—*RBM* and *TSPY*—may also be the result of the Y chromosome having acquired and amplified autosomal genes (*26*). We speculate that the selective advantage conferred by the NRY's retention and amplification of male fertility factors (from throughout the genome) accounts for the multitude of testis-specific gene families there. These activities may have been preeminent in shaping the NRY's gene repertoire, because it appears that most NRY transcription units are members of testis-specific families (*27*). We suspect that most of the NRY's transcription units do not date from the Y chromosome's common ancestry with the X chromosome but instead are more recent acquisitions.

The importance of the human Y chromosome in fertility has been underscored by recent genetic studies. Many men with spermatogenic failure, although otherwise healthy, lack portions of the NRY (*7*, *8*, *28*). These findings have suggested the existence of NRY genes that play critical roles in male germ cell development but are not required elsewhere in the body. Previous deletion-mapping studies have implicated four regions of the NRY in either spermatogenic failure or germ cell tumorigenesis, and in each of the four regions we now report novel candidate genes expressed specifically, or most abundantly (*29*), in testes (Figs. 1 and 2).

Although X-homologous and testis-specific genes are somewhat intermingled within the NRY, clustering is evident (Fig. 1). The geographic distribution of the two classes correlates well with previously identified sequence domains within the euchromatic NRY (*10*, *11*). Ten of the 11 known testis-specific families map to previously identified regions of Y-specific repetitive sequences (*30*). Indeed, one or more testis-specific gene families are found in nearly all known regions of euchromatic Y repeats (Fig. 1). Ironically, it had been widely assumed, partly on theoretical grounds, that these domains consisted of "junk" DNA (*12*). To the contrary, our results argue that these Y-specific repetitive regions are gene-rich, containing most of the NRY's transcription units (*27*). We speculate that these were regions of rampant gene amplification during mammalian evolution. By contrast, none of the eight X-homologous genes map to the Y-repeat domains; they all map to regions previously identified as consisting largely of single-copy (or in some cases X-homologous) sequences. We postulate that, earlier in mammalian evolution, these regions of the NRY shared extensive nucleotide sequence identity with the X chromosome.

Although more genes probably remain to be discovered, the 20 genes and families shown in Fig. 1 may constitute the majority of NRY genes, and full-length cDNA sequences are available for 18 of them. The stage is now set for systematic evolutionary, biochemical, and cell-biological studies of this distinctive segment of the human genome.

REFERENCES AND NOTES

1. The remaining 5% of the chromosome is composed of two pseudoautosomal regions that maintain sequence identity with the X chromosome by meiotic recombination [H. J. Cooke, W. R. Brown, G. A. Rappold, *Nature* **317**, 687 (1985); M. C. Simmler *et al.*, *ibid.*, p. 692; D. Freije, C. Helms, M. S. Watson, H. Donis-Keller, *Science* **258**, 1784 (1992); N. A. Affara *et al.*, *Cytogenet. Cell Genet.* **73**, 33 (1996); E. Rao *et al.*, *Nature Genet.* **16**, 54 (1997)].
2. J. J. Bull, *Evolution of Sex Determining Mechanisms* (Benjamin Cummings, Menlo Park, CA, 1983); J. A. Graves, *Annu. Rev. Genet.* **30**, 233 (1996); B. Charlesworth, *Curr. Biol.* **6**, 149 (1996); W. R. Rice, *BioScience* **46**, 331 (1996).
3. C. Stern, *Am. J. Hum. Genet.* **9**, 147 (1957).
4. P. A. Jacobs and J. A. Strong, *Nature* **183**, 302 (1959); C. E. Ford, K. W. Jones, P. E. Polani, J. C. Almida, J. H. Briggs, *Lancet* **i**, 711 (1959).
5. D. C. Page *et al.*, *Cell* **51**, 1091 (1987); A. H. Sinclair *et al.*, *Nature* **346**, 240 (1990); J. Arnemann, S. Jakubiczka, S. Thuring, J. Schmidtke, *Genomics* **11**, 108 (1991); E. C. Salido, P. H. Yen, K. Koprivnikar, L. C. Yu, L. J. Shapiro, *Am. J. Hum. Genet.* **50**, 303 (1992); A. I. Agulnik, M. J. Mitchell, J. L. Lerner, D. R. Woods, C. E. Bishop, *Hum. Mol. Genet.* **3**, 873 (1994).
6. E. M. Fisher *et al.*, *Cell* **63**, 1205 (1990).
7. K. Ma *et al.*, *ibid.* **75**, 1287 (1993).
8. R. Reijo *et al.*, *Nature Genet.* **10**, 383 (1995).
9. J. A. Graves, *Bioessays* **17**, 311 (1995).
10. D. Vollrath *et al.*, *Science* **258**, 52 (1992).
11. S. Foote, D. Vollrath, A. Hilton, D. C. Page, *ibid.*, p. 60.
12. B. Charlesworth, *ibid.* **251**, 1030 (1991); E. Seboun *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **1**, 237 (1986).
13. We performed cDNA selection [M. Lovett, J. Kere, L. M. Hinton, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9628 (1991)] with bulk cDNA from human adult testes (Clontech, Palo Alto, CA) and, as selector, a cosmid library prepared from flow-sorted Y chromosomes (Lawrence Livermore National Laboratory: LL0YNC03). Thirty-six hundred random cosmids, providing nearly fivefold coverage of the 30-Mb euchromatic region, were used to generate 150 pools of selector DNA. (Theoretically, these 3600 cosmids are expected to contain 99% of all euchromatic Y sequences. It is our experience that this cosmid library contains >95% of Y-DNA sequence-tagged sites tested.) With each selector pool, we carried out four successive rounds of cDNA selection, followed by two rounds of subtraction with human COT-1 DNA (Gibco-BRL, Gaithersburg, MD) to remove highly repetitive sequences. A plasmid library was prepared from each of the 150 resulting pools of selected cDNA fragments, and 24 clones from each library were sequenced from one end. Of the 3600 sequences generated, about 600 were of poor technical quality and about 500 were found to derive from the cloning vector or the *E. coli* host, leaving 2539 sequences for further analysis. Of the 2539 sequence fragments, 536 corresponded to previously reported NRY genes (487 to *TSPY*, 15 to *RBM*, 14 to *RPS4Y*, 9 to *SMCY*, 5 to *DAZ*, 3 to *SRY*, and 3 to *ZFY*), and 43 corresponded to previously reported pseudoautosomal genes (15 to *XE7*, 11 to *CSF2RA*, 4 to *IL3RA*, 3 to *ASMT*, 3 to *IL9R*, 2 to *ANT3*, 2 to *MIC2*, 2 to *SHOX*, and 1 to *SYBL1*). Electronic analysis of the roughly 2000 remaining sequences revealed that about 200 contained known repetitive elements, and these were not pursued. Through electronic identification of redundancies and sequence overlaps, the remaining sequences were reduced to 1093 sequence contigs. Sequences representing these 1093 contigs were individually hybridized to dot-blotted yeast genomic DNAs of 60 yeast artificial chromosomes (YACs) comprising most of the Y's euchromatic region (*11*). One hundred eighty-one sequences that hybridized to the great majority of the YACs were judged likely to contain highly repeated elements and were not pursued, leaving 912 sequences for further analysis. These sequences were individually hybridized to Southern blots of Eco RI–digested human 46,XX female and 49,XYYYY male genomic DNAs; hybridization was done at 65°C in 0.5 M Na_iPO_4 (pH 7.5) and 7% SDS; washing was at 65°C in 1× saline sodium citrate (SSC) and 0.1% SDS. Eight hundred thirty-two hybridizations yielded interpretable results. Many sequences appeared to contain highly repeated elements common to males and females or failed to detect an unambiguously Y-specific restriction fragment, and these were not pursued. (This eliminated sequences derived from the pseudoautosomal regions or other regions of extremely high X-Y nucleotide similarity.) By contrast, 308 sequences hybridized to at least one prominent fragment that was present in 49,XYYYY but absent in 46,XX, which suggests that these sequences derived from the NRY. Each of these 308 sequences was individually used to screen, by hybridization, about 2 million plaques from a λ phage library of human adult testis cDNA (Clontech).
14. Genes were localized on a previously reported NRY deletion map by polymerase chain reaction (PCR) testing of individuals carrying partial Y chromosomes (*10*). Most genes were localized to a single deletion interval. Some genes could not be unambiguously placed, because copies exist in multiple locations in the NRY. In such cases, genes were localized by PCR testing of YACs encompassing the euchromatic region (*11*). Homologs of *DBY*, *TB4Y*, *EIF1AY*, and *UTY* were mapped to the human X chromosome by PCR testing of a panel of human/rodent somatic hybrid cell lines (Research Genetics, Huntsville, AL). PCR conditions and primer sequences have been deposited at GenBank, where accession numbers are as follows: *DBY*, G34990; *TB4Y*, G34981; *EIF1AY*, G34991; *UTY*, G34977; *DFFRY*, G34983; *CDY*, G34975; *BPY1*, G34985; *BPY2*, G34986; *XKRY*, G34987; *PRY*, G34984; *TTY1*, G34978; *TTY2*, G34980; *DBX*, G34988; *TB4X*, G34979; *EIF1AX*, G34989; *UTX*, G34976; and *DFFRX*, G34982. *TB4X* primers were designed from an unreported intron; all other PCR primers were chosen from cDNA sequences.
15. S. Mumm, B. Molini, J. Terrell, A. Srivastava, D. Schlessinger, *Genome Res.* **7**, 307 (1997); A. Schwartz *et al.*, *Hum. Mol. Genet.*, in press.
16. Two previously identified NRY genes do not fit neatly into either of the two classes: *SRY* is expressed in testes but exists in only one copy in the NRY. *AMELY* and its X-linked homolog *AMELX* are expressed only in tooth buds. The X inactivation status of *AMELX* is unknown. Thus, not all testis-specific NRY genes are multicopy, and not all X-homologous NRY genes are ubiquitously expressed.
17. H. Gondo *et al.*, *J. Immunol.* **139**, 3840 (1987); D. Safer, M. Elzinga, V. T. Nachmias, *J. Biol. Chem.* **266**, 4029 (1991).
18. T. E. Dever *et al.*, *J. Biol. Chem.* **269**, 3212 (1994); J. W. Hershey, *Annu. Rev. Biochem.* **60**, 717 (1991).
19. M. H. Jones *et al.*, *Hum. Mol. Genet*. **5**, 1695 (1996).
20. A. Schneider-Gädicke, P. Beer-Romero, L. G. Brown, R. Nussbaum, D. C. Page, *Cell* **57**, 1247 (1989); A. I. Agulnik *et al.*, *Hum. Mol. Genet.* **3**, 879 (1994).
21. For each of the previously untested X-linked genes *DBX*, *TB4X*, and *EIF1AX*, we assayed X inactivation status by two methods: (i) reverse transcription (RT)–PCR on human-rodent hybrids retaining inactive human X chromosomes and (ii) CpG methylation studies in which male and female genomic DNAs digested with methylation-sensitive restriction endonucleases were used as templates for PCR [S.-W. Luoh *et al.*, *Genomics* **29**, 353 (1995)]. The previously untested gene *UTX* was assayed by the first method only. As judged by these assays, each of the four genes escapes X inactivation (B. Lahn and D. C. Page, unpublished results).
22. M. Watanabe, A. R. Zinn, D. C. Page, T. Nishimoto, *Nature Genet.* **4**, 268 (1993).
23. M. A. Ferguson-Smith, *J. Med. Genet.* **2**, 142 (1965); A. R. Zinn, D. C. Page, E. M. Fisher, *Trends Genet.* **9**, 90 (1993).
24. The potential role of *RPS4Y* and *RPS4X* in Turner syndrome is controversial (*6*) [W. Just, C. Geerkens, K. R. Held, W. Vogel, *Hum. Genet.* **89**, 240 (1992)]. At least one Turner gene maps to the Xp-Yp pseudoautosomal region (*1*).
25. R. A. Fisher, *Biol. Rev.* **6**, 345 (1931).
26. R. Saxena *et al.*, *Nature Genet.* **14**, 292 (1996); M. L. Delbridge *et al.*, *ibid.* **15**, 131 (1997).
27. Each NRY testis-specific gene family has multiple members—20 to 40 copies in the case of *TSPY* [E. Manz, F. Schnieders, A. M. Brechlin, J. Schmidtke, *Genomics* **17**, 726 (1993)] and 20 or more copies in the case of *RBM* (*7*). All together, the various Y-specific gene families may include several hundred genes or copies. Though it is not known how many of these are functional, it seems likely that Y-specific, testis-specific gene families make up most NRY transcription units.
28. P. H. Vogt *et al.*, *Hum. Mol. Genet.* **5**, 933 (1996); J. L. Pryor *et al.*, *N. Engl. J. Med.* **336**, 534 (1997).
29. Two regions implicated in spermatogenic failure (Fig. 1) contain X-homologous genes *DBY* and *EIF1AY*, with abundant testis-specific transcripts in addition to ubiquitous transcripts (Fig. 2). Expression of genes important in male germ cell development is not necessarily restricted to the testis.
30. The single exception is *BPY1*, which cross-hybridizes to the X chromosome and maps to a previously recognized region of X homology (*11*).
31. K. Tsuchiya, R. Reijo, D. C. Page, C. M. Disteche, *Am. J. Hum. Genet.* **57**, 1400 (1995); P. Salo *et al.*, *Genes Chromosomes Cancer* **14**, 210 (1995); P. Salo, H. Kaariainen, D. C. Page, A. de la Chapelle, *Hum. Genet.* **95**, 283 (1995).
32. P. Linder *et al.*, *Nature* **337**, 121 (1989); R.-Y. Chuang, P. L. Weaver, Z. Liu, T.-H. Chang, *Science* **275**, 1468 (1997).
33. A. Greenfield *et al.*, *Nature Genet.* **14**, 474 (1996); R. S. Sikorski, M. S. Boguski, M. Goebl, P. Hieter, *Cell* **60**, 307 (1990); D. Tzamarias and K. Struhl, *Genes Dev.* **9**, 821 (1995).
34. J. A. Fischer-Vize, G. M. Rubin, R. Lehmann, *Development* **116**, 985 (1992); Y. Huang, R. T. Baker, J. A.

Fischer-Vize, *Science* **270**, 1828 (1995).
35. T. C. James and S. C. Elgin, *Mol. Cell Biol*. **6**, 3862 (1986); R. Paro and D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 263 (1991); B. Tschiersch *et al.*, *EMBO J.* **13**, 3822 (1994); M. T. Madireddi *et al.*, *Cell* **87**, 75 (1996); D. G. Stokes, K. D. Tartof, R. P. Perry, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7137 (1996).
36. P. M. Palosaari *et al.*, *J. Biol. Chem.* **266**, 10750 (1991); A. Schmitz, K. H. Gartemann, J. Fiedler, E. Grund, R. Eichenlaub, *Appl. Environ. Microbiol.* **58**, 4068 (1992); V. Sharma, K. Suvarna, R. Meganathan, M. E. Hudspeth, *J. Bacteriol.* **174**, 5057 (1992); M. Kanazawa *et al.*, *Enzyme Protein* **47**, 9 (1993); Z. L. Boynton, G. N. Bennet, F. B. Rudolph, *J. Bacteriol.* **178**, 3015 (1996).
37. M. Ho *et al.*, *Cell* **77**, 869 (1994).
38. W. Hendriks *et al.*, *J. Cell Biochem.* **59**, 418 (1995).
39. We thank H. Skaletsky and F. Lewitter for help with sequence analysis; Lawrence Livermore National Laboratory for the flow-sorted Y cosmid library; and P. Bain, A. Bortvin, A. de la Chapelle, G. Fink, K. Jegalian, T. Kawaguchi, E. Lander, H. Lodish, P. Matsudaira, D. Menke, U. RajBhandary, R. Reijo, S. Rozen, A. Schwartz, C. Sun, and C. Tilford for comments on the manuscript. Supported by NIH.

28 April 1997; accepted 9 September 1997

Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale

Joseph L. DeRisi, Vishwanath R. Iyer, Patrick O. Brown*

DNA microarrays containing virtually every gene of *Saccharomyces cerevisiae* were used to carry out a comprehensive investigation of the temporal program of gene expression accompanying the metabolic shift from fermentation to respiration. The expression profiles observed for genes with known metabolic functions pointed to features of the metabolic reprogramming that occur during the diauxic shift, and the expression patterns of many previously uncharacterized genes provided clues to their possible functions. The same DNA microarrays were also used to identify genes whose expression was affected by deletion of the transcriptional co-repressor *TUP1* or overexpression of the transcriptional activator *YAP1*. These results demonstrate the feasibility and utility of this approach to genomewide exploration of gene expression patterns.

The complete sequences of nearly a dozen microbial genomes are known, and in the next several years we expect to know the complete genome sequences of several metazoans, including the human genome. Defining the role of each gene in these genomes will be a formidable task, and understanding how the genome functions as a whole in the complex natural history of a living organism presents an even greater challenge.

Knowing when and where a gene is expressed often provides a strong clue as to its biological role. Conversely, the pattern of genes expressed in a cell can provide detailed information about its state. Although regulation of protein abundance in a cell is by no means accomplished solely by regulation of mRNA, virtually all differences in cell type or state are correlated with changes in the mRNA levels of many genes. This is fortuitous because the only specific reagent required to measure the abundance of the mRNA for a specific gene is a cDNA sequence. DNA microarrays, consisting of thousands of individual gene sequences printed in a high-density array on a glass microscope slide (*1, 2*), provide a practical and economical tool for studying gene expression on a very large scale (*3–6*).

Saccharomyces cerevisiae is an especially favorable organism in which to conduct a systematic investigation of gene expression. The genes are easy to recognize in the genome sequence, *cis* regulatory elements are generally compact and close to the transcription units, much is already known about its genetic regulatory mechanisms, and a powerful set of tools is available for its analysis.

A recurring cycle in the natural history of yeast involves a shift from anaerobic (fermentation) to aerobic (respiration) metabolism. Inoculation of yeast into a medium rich in sugar is followed by rapid growth fueled by fermentation, with the production of ethanol. When the fermentable sugar is exhausted, the yeast cells turn to ethanol as a carbon source for aerobic growth. This switch from anaerobic growth to aerobic respiration upon depletion of glucose, referred to as the diauxic shift, is correlated with widespread changes in the expression of genes involved in fundamental cellular processes such as carbon metabolism, protein synthesis, and carbohydrate storage (*7*). We used DNA microarrays to characterize the changes in gene expression that take place during this process for nearly the entire genome, and to investigate the genetic circuitry that regulates and executes this program.

Yeast open reading frames (ORFs) were amplified by the polymerase chain reaction (PCR), with a commercially available set of primer pairs (*8*). DNA microarrays, containing approximately 6400 distinct DNA sequences, were printed onto glass slides by using a simple robotic printing device (*9*). Cells from an exponentially growing culture of yeast were inoculated into fresh medium and grown at 30°C for 21 hours. After an initial 9 hours of growth, samples were harvested at seven successive 2-hour intervals, and mRNA was isolated (*10*). Fluorescently labeled cDNA was prepared by reverse transcription in the presence of Cy3(green)- or Cy5(red)-labeled deoxyuridine triphosphate (dUTP) (*11*) and then hybridized to the microarrays (*12*). To maximize the reliability with which changes in expression levels could be discerned, we labeled cDNA prepared from cells at each successive time point with Cy5, then mixed it with a Cy3-labeled "reference" cDNA sample prepared from cells harvested at the first interval after inoculation. In this experimental design, the relative fluorescence intensity measured for the Cy3 and Cy5 fluors at each array element provides a reliable measure of the relative abundance of the corresponding mRNA in the two cell populations (Fig. 1). Data from the series of seven samples (Fig. 2), consisting of more than 43,000 expression-ratio measurements, were organized into a database to facilitate efficient exploration and analysis of the results. This database is publicly available on the Internet (*13*).

During exponential growth in glucose-rich medium, the global pattern of gene expression was remarkably stable. Indeed, when gene expression patterns between the first two cell samples (harvested at a 2-hour interval) were compared, mRNA levels differed by a factor of 2 or more for only 19 genes (0.3%), and the largest of these differences was only 2.7-fold (*14*). However, as glucose was progressively depleted from the growth media during the course of the experiment, a marked change was seen in the global pattern of gene expression. mRNA levels for approximately 710 genes were induced by a factor of at least 2, and the mRNA levels for approximately 1030 genes declined by a factor of at least 2. Messenger RNA levels for 183 genes increased by a factor of at least 4, and mRNA levels for 203 genes diminished by a factor of at least 4. About half of these differentially expressed genes have no currently recognized function and are not yet named. Indeed, more than 400 of the differentially expressed genes have no apparent homology

Department of Biochemistry, Stanford University School of Medicine, Howard Hughes Medical Institute, Stanford, CA 94305–5428, USA.

*To whom correspondence should be addressed. E-mail: pbrown@cmgm.stanford.edu