A primitive Y chromosome in papaya marks incipient sex chromosome evolution

Zhiyong Liu[1], Paul H. Moore[2], Hao Ma[1,3], Christine M. Ackerman[1], Makandar Ragiba[1], Qingyi Yu[1,3], Heather M. Pearl[1], Minna S. Kim[1], Joseph W. Charlton[1], John I. Stiles[4], Francis T. Zee[2], Andrew H. Paterson[5] & Ray Ming[1]

[1]*Hawaii Agriculture Research Center, Aiea, Hawaii 96701, USA*
[2]*USDA-ARS, Pacific Basin Agricultural Research Center, Hilo, Hawaii 96720, USA*
[3]*Department of Molecular Biosciences and Bioengineering, University of Hawaii, Honolulu, Hawaii 96822, USA*
[4]*Integrated Coffee Technologies Incorporated, Waialua, Hawaii 96791, USA*
[5]*Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602, USA*

Many diverse systems for sex determination have evolved in plants and animals[1–3]. One involves physically distinct (heteromorphic) sex chromosomes (X and Y, or Z and W) that are homozygous in one sex (usually female) and heterozygous in the other (usually male). Sex chromosome evolution is thought to involve suppression of recombination around the sex determination genes, rendering permanently heterozygous a chromosomal region that may then accumulate deleterious recessive mutations by Muller's ratchet, and fix deleterious mutations by hitchhiking as nearby favourable mutations are selected on the Y chromosome[4,5]. Over time, these processes may cause the Y chromosome to degenerate and to diverge from the X chromosome over much of its length; for example, only 5% of the human Y chromosome still shows X–Y recombination[6]. Here we show that papaya contains a primitive Y chromosome, with a male-specific region that accounts for only about 10% of the chromosome but has undergone severe recombination suppression and DNA sequence degeneration. This finding provides direct evidence for the origin of sex chromosomes from autosomes.

Papaya (*Carica papaya* L., $2n = 18$), a polygamous angiosperm with male, female and hermaphroditic forms, offers several advantages for genetic and evolutionary studies. These include a small genome of 372 megabases (Mb), a short generation time of 9–15 months, numerous flower types, a unique evolutionary process in female flowers (see Supplementary Information), an intriguing system of sex determination and an established transformation system. Sex determination in papaya has been a frequent subject of genetic analyses[7–10] because it is directly related to efficient commercial fruit production.

Hofmeyr[7] and Storey[8] each concluded that sex in papaya is determined by a single gene with three alleles: M, male; M^h, hermaphrodite; and m, female. Females (mm) are homozygous recessive. Males (Mm) and hermaphrodites (M^hm) are enforced sex heterozygotes, with typically 25% of seeds nonviable in fruit, because every combination of dominant alleles (MM, MM^h or M^hM^h) is embryonic lethal. Hofmeyr[9] later suggested that M and M^h represent genetically inactive regions of the sex chromosomes from which vital genes are missing. In Storey's[10] final hypothesis, the sex-determining region includes a complex of genes regulating stamen and carpel development, a lethal factor and a recombination-suppressing factor. Because males and hermaphrodites are heterogametic, whereas females are homogametic, sex determination in papaya has been considered by some to be of the XY chromosome type even though heteromorphic chromosomes have not been found[1,11].

The papaya sex locus has been genetically mapped to linkage group 1 (LG 1; ref. 12). In a high-density genetic map containing 1,501 amplified fragment length polymorphism (AFLP) and morphological markers from 54 F_2 plants, a total of 225 markers (that is, 15% of all markers mapped on the genome and 66% of the 342 markers on LG 1) co-segregate with the sex locus[13]. The remaining 117 AFLP markers on LG 1 detect regions of active recombination spanning 289.7 cM. These data reveal an extremely high level of DNA polymorphism in the genomic region immediately surrounding the sex locus.

We fine-mapped the sex determination gene (Table 1 and Supplementary Fig. 1), using 4,380 informative chromosomes (two each from 2,190 female and hermaphrodite plants from three F_2 and one F_3 populations), and two sequence-characterized amplified region (SCAR) markers (W11 and T12)[14], three cloned sex-linked AFLP markers (cpsm10, cpsm31 and cpsm54) and one cloned bacterial artificial chromosome (BAC) end (cpbe55). No recombinants were detected. The finding that a region so recombinationally small contains such a large percentage of polymorphic markers indicates that the two homologues are highly differentiated in this region. This is consistent with the classical notion that an early stage of sex chromosome evolution involves suppression of recombination around the sex determination locus, leading to gradual degeneration of the Y chromosome.

We physically mapped the non-recombining region by using a 13.7 × BAC library[15]. The sex-linked SCAR marker W11 identified four positive BACs. A BAC contig assembled by chromosome walking spanned 990 kilobases (kb). Forty-two cpsm markers and three previously available sex co-segregating markers, T12, PSDM and Nafp (see Supplementary Information), hybridized to additional BACs. Among these 42 cpsm markers, 24 (57%) could be placed on BAC contig maps, 3 identified individual contigs (data not shown), and 15 contained repetitive sequences (and thus could not be mapped). We cloned (iteratively) 92 BAC ends from contig-terminal BACs and used them to close the gaps. All relevant BAC clones were fingerprinted and mapped to verify the hybridization-based physical map. The 2.5-Mb physical map includes two major and three minor contigs, containing 4 SCAR, 82 cpbe and 24 cpsm loci (Fig. 1).

Although gaps still remain in the physical map, at least 57% of the cpsm markers are located in a small (2.5-Mb) region of the sex-determining chromosome. On the basis of the papaya's genome size and the cytological observation that its chromosomes are similar in size[16], each papaya chromosome is estimated to be at about 41 Mb. With a mapping population of 54 plants[13] and the use of dominant markers (that is, only one gamete is genetically informative), the absence of recombinants suggests that two loci must be less than about 4 cM apart with 95% confidence[17], which delimits the maximum size of the male-specific region (MSY). The lack of recombinants in 2,190 plants suggests that the 225 sex

Table 1 **Results of fine-mapping with SCAR markers in the MSY region**

Population	Progeny	Hermaphrodite	Female	SCAR markers	Recombinant
Kapoho × SunUp	F_2	335	150	W11	0
Kapoho × SunUp	F_2	335	156	W11,T12, cpbe55	0
Kapoho × SunUp	F_3	481	274	W11,T12, cpbe55	0
Kapoho × Saipan Red	F_2	175	49	cpsm31, cpsm10	0
AU9 × SunUp	F_2	170	65	W11,T12, cpsm54	0
Total	2,190	1,496	694		0

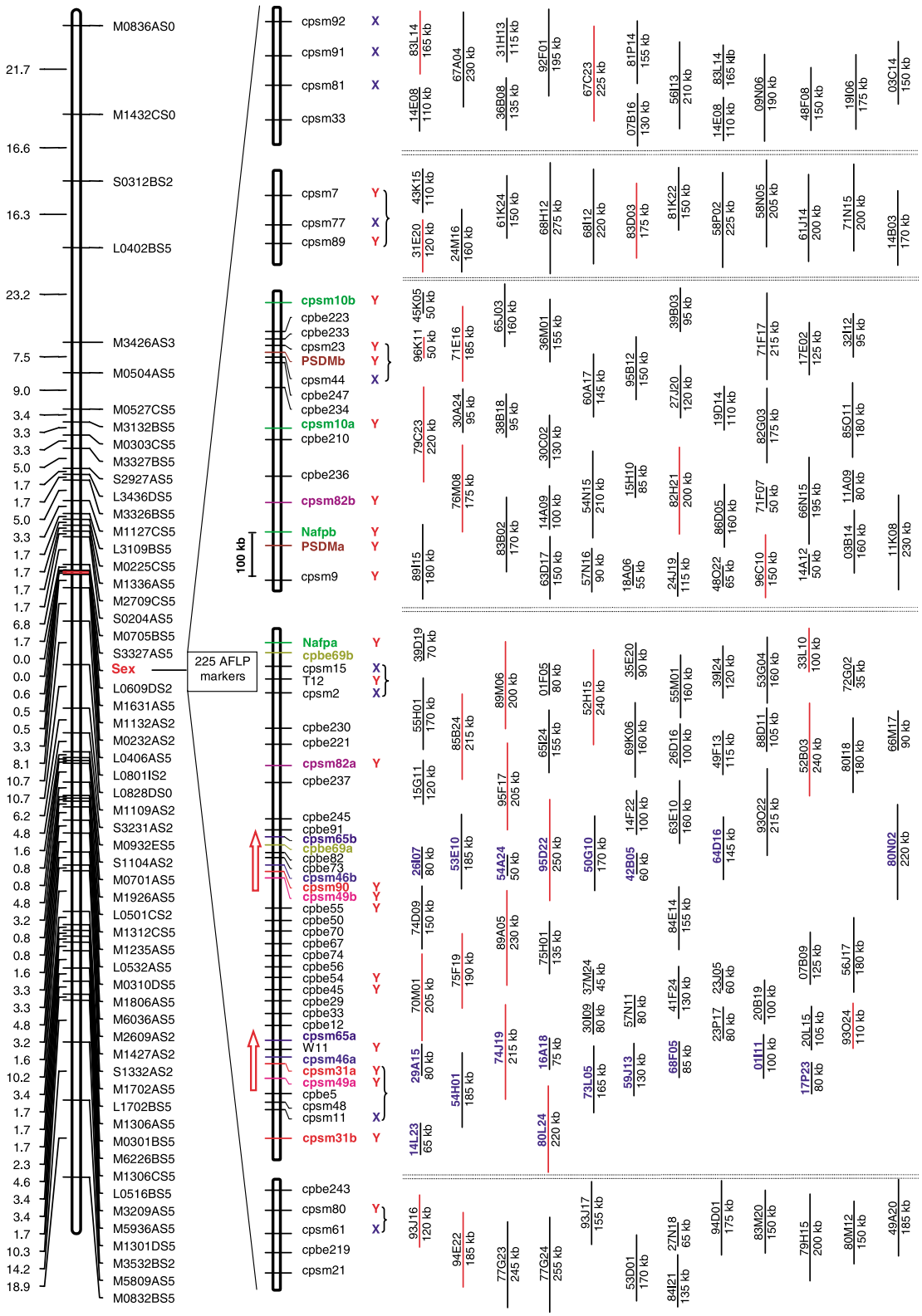

Figure 1 AFLP map of papaya LG 1 and physical maps of BAC contigs in the MSY region. Left, AFLP map of LG 1. Right, physical maps of two long and three short BAC contigs in the MSY region. Arrows indicate a large tandem duplication; marker names shown in the same colour indicate additional smaller duplications. The letters X (blue) and Y (red) indicate whether the sequence is X-like or Y-like. The X- and Y-like DNA sequences located within one BAC are delineated by a bracket. BAC clones used for sequencing from the MSY are shown in red. BAC clones tested in Fig. 2c are marked in blue.

co-segregating markers are located in a region that is much less than 4 cM. The remaining 117 markers on LG 1 span 289.7 cM, however, indicating that much of this chromosome is still recombining with its homologue[13]. Among the markers used for fine-mapping, W11 and T12 are about 900 kb apart, as estimated by adding the insert sizes of non-overlapping BAC clones. On the basis of the total genetic map of 3,294 cM and the genome size of 372 Mb, this is equivalent to 8 cM in other genomic regions with average recombination rates. Thus, the 1,481 plants informative in fine-mapping with both W11 and T12 would have been expected to contain 237 recombinant plants, but none was found.

A mosaic arrangement of conserved (X-like) and diverged (Y-like) sequences (Fig. 1) shows that degradation of the Y chromosome is distributed across the non-recombining region. Out of 26 test sequences designed from the physically mapped non-recombining region in male, female and hermaphrodite plants (see Methods), 17 were Y-like (Fig. 1, 'Y') and present in hermaphrodite and male but not female, whereas 9 were X-like (Fig. 1, 'X') and present in all three sex types. Six X-like sequences were on BACs that also contained Y-like sequences (Fig. 1, brackets).

Duplicated sequences are common in the non-recombining region. Out of 109 DNA markers on the physical map, 9 identified duplications, including 5 (31%) Y-like test sequences. A tandem duplication identified by cpsm31, cpsm46, cpsm49, cpsm65 and cpsm90 spanned more than 100 kb, with the repeats located 500 kb apart (Fig. 1). Nucleotide insertions, deletions and substitutions were observed in all duplicated DNA sequences such as cpsm31 and cpsm90 (Fig. 2 and Supplementary Fig. 2).

Random subclone sequences (628 reads, totalling 513 kb) from 25 non-redundant BACs of the non-recombining region were analysed to determine the nature of the sequences and to estimate the gene density. We found that 48 reads from 44 subclones matched functional genes, suggesting an average density of 8.6 genes per 100 kb. Using RepeatMasker (http://repeatmasker.genome.washington.edu) to search against the *Arabidopsis thaliana* repeat database, we identified 95 long terminal repeat (LTR) retroelements (including 82 Gypsy-like and 13 Copia-like elements), with an average of 18.5 LTRs per 100 kb. Forty short inverted repeats were identified (7.8 per 100 kb). In comparison to a genome-wide sample of papaya DNA based on 684 reads totalling 517 kb (GenBank Accessions CG026197 to CG026996), the MSY region showed 37.7% lower gene density, 27.6% higher retroelement density, and 188.9% higher inverted repeat density.

The severe suppression of recombination and extensive divergence between homologues in the region containing the papaya sex-determining genes indicates that this is an incipient sex chromosome. On the basis of the size of the present contig map (2.5 Mb) and the 57% of cpsm markers that have been accounted for, the physical size of the MSY is estimated at 4–5 Mb or 10% of papaya's primitive Y chromosome.

The discovery of an incipient Y chromosome in papaya, of which 10% is a non-recombining, rapidly evolving, sex-determining region flanked by normal autosome-like regions that comprise the remaining 90% of the chromosome (Fig. 3), provides direct evidence for the theory that sex chromosomes evolve from autosomes[3]. It also supports the hypothesis that sex chromosomes arise from suppression of recombination around the sex-determining locus[1,3,18]. The papaya MSY region may resemble the ancestor of the human Y chromosome as it existed about 240–320 million years ago, when recombination suppression is postulated to have been limited to only a small portion of the chromosome[19]. Analysis of the pseudoautosomal boundary in mammals indicates that sex chromosome differentiation in humans probably started near male-determining loci by suppression of homologous recombination[20].

The incipient sex chromosomes of papaya may yield insights about earlier stages of sex chromosome evolution. The small physical size of the MSY region and the mosaic arrangement of sequence degradation indicate a recent origin of the papaya sex chromosomes (see Supplementary Information). By contrast, in the better-studied X and Y chromosomes of the plant *Silene latifolia*, 90% of the Y chromosome, albeit mostly euchromatic DNA, seems to be degenerated[1,21,22]. Although heteromorphic sex chromosomes have not been detected in papaya, a high frequency of precocious separation of one chromosome pair during meiosis has been reported in the anthers of male and hermaphrodite plants but not in the pistils of female plants (ref. 24 and Supplementary Fig. 3), which is compatible with relaxed pairing between primitive Y and X chromosomes owing to the existence of a non-pairing MSY region.

The lethal effect of homozygous dominant sex-determining alleles in papaya[7–10] provides further evidence of the degeneration of the MSY region. Because papaya seeds with the homozygous Y chromosomes are aborted 25–50 d after pollination[23], genes that are

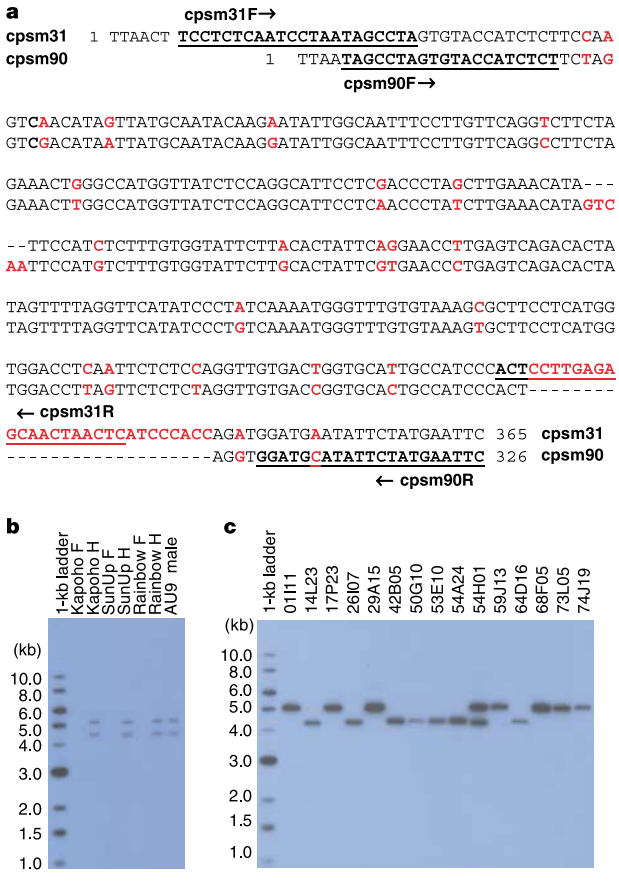

Figure 2 Characterization of duplicated markers cpsm31 and cpsm90 in the MSY region. **a**, DNA sequence comparison of cpsm31 and cpsm90. PCR primer sequences are underlined. Nucleotide insertion, deletion and transition or transversion are indicated in red. **b**, Southern hybridization pattern of cpsm31 on *Hin*dIII-digested genomic DNA of different sex types. Two cpsm31 loci are present in hermaphrodites (H) and males, but are absent from females (F). **c**, Southern hybridization of cpsm31 on *Hin*dIII-digested BAC DNA from the tandemly duplicated region of MSY.

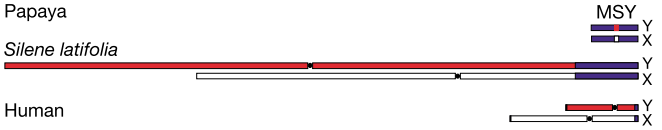

Figure 3 Comparative organization of Y chromosomes in papaya (estimated as 41 Mb), *S. latifolia* (estimated as 570 Mb) and human (66 Mb). The extent of MSY regions (red) and recombinationally active regions (blue) is shown. Filled circles represent the centromeres. The location of the centromere on the papaya Y chromosome remains unknown.

vital for papaya embryo development may once have existed in what is now the MSY region. Future detailed comparisons of the MSY and its homologue, as well as related chromosomes from other taxa that do not have discrete sexes, may help to identify these genes.

DNA sequence duplication and repetition in the MSY regions of papaya, *Drosophila* and humans may reflect common processes associated with evolution of the sex chromosomes in plants and animals. Complete sequencing of the euchromatic DNA on the human Y chromosome has identified X-transposed, X-degenerated and ampliconic sequences[6]. X-degenerated sequences seem to be common in the papaya MSY, because 65% of the test sequences were diverged between the primitive X and Y chromosomes and about a third of the MSY is degenerated on the basis of gene density. Ampliconic sequences are probably present on the papaya primitive Y chromosome, because abundant inverted repeats were found and 38% of random DNA markers in the MSY detected repetitive sequences. Sequence duplication and retroelement accumulation may be driving the degeneration of the papaya MSY region. X- and Y-linked homologues in *Silene* also show deletion and accumulation of repetitive sequences, as well as low variability in a Y-linked gene[22]. The liverwort Y chromosome is a mosaic structure of repeats and subrepeats[25]. In *Drosophila miranda*, genes show various signs of degeneration on the non-recombining neo-Y, and repetitive DNA is accumulated[26].

We found that hermaphrodite and male papaya plants share identical DNA sequences in most parts of the MSY region (Supplementary Fig. 2 and Table S1). These two sex types seem to share a haplotype for this region that differs from that of females and is recently derived from a common ancestral chromosome. This observation suggests that the hermaphrodites with an MSY region are already genetically different from the ancestral hermaphrodites. The findings in papaya strongly support the hypothesis that Y chromosome evolution involved at least two steps. First, a single male-sterility mutation led to femaleness and to selection for absence of recombination; second, the Y chromosome evolved as hermaphrodites converted to males by a second mutation or mutations led to female sterility[18]. Thus, gynodioecy might be an intermediate step to dioecy in the family Caricaceae[27].

Papaya provides an extraordinary opportunity to elucidate the evolutionary process of dioecy because monoecious, dioecious and perfect flowers all exist among its close relatives. Only two species, *C. papaya* and *Vasconcella cundinamarcensis*, are polygamous with all three sex types; *V. monoica* is monoecious; and the other 32 species in the family Caricaceae are dioecious. DNA sequencing will allow comparative analyses of the MSY region on the primitive Y chromosome of monoecious and dioecious species of Caricaceae to shed light on the evolution of sex chromosomes and dioecy.

Although higher plants can regulate sex by either autosomal genes or sex chromosomes[21], few flowering plants have evolved heteromorphic sex chromosomes analogous to those of many animals. Despite the recent origin of plant sex chromosomes, they share several characteristics with those in animals, such as prevalence of male heterogamy, degeneration of the Y chromosome, and lack of recombination between X and Y chromosomes. These parallel properties indicate that many similar, if not identical, evolutionary forces may drive sex chromosome evolution in plants and animals. The primitive sex chromosomes in papaya provide a unique opportunity to test hypotheses and theories about sex chromosome evolution near the inception of this process[18]. □

Methods

Plant materials

Female and hermaphrodite plants of the Hawaiian Solo type papaya cultivars Kapoho and SunUp and their F_1 progeny Rainbow, as well as male plants of the Australian variety AU9, were used for polymerase chain reaction (PCR) and Southern hybridization analyses. We used three F_2 populations derived from single self-pollinated hermaphrodite F_1 plants from each of the three crosses: Kapoho × SunUp, Kapoho × Saipan Red and AU9 × SunUp. The F_3 population of Kapoho × SunUp was from self-pollinated hermaphrodite F_2 plants.

DNA markers used for physical mapping

Three kinds of sex-linked DNA marker were used to screen the papaya BAC library and to construct the BAC contigs: SCAR markers from previous research[14], *Carica papaya* BAC ends (cpbe) cloned by a PCR approach[28], and cloned sex-linked AFLP markers (cpsm) identified from a papaya high-density map[13]. We excised the sex-linked AFLP fragments from manual sequencing gels after silver staining and cloned them into a pCRII-TOPO vector (Invitrogen). DNA sequencing was done on a DNA sequencer (Li-Cor) by using infrared dye-labelled primers. PCR primers were designed for each cpsm and were used to amplify genomic DNA from females and hermaphrodites of the cultivars Kapoho, SunUp and Rainbow, and from a male of AU9.

Contig map construction

We screened our papaya BAC library with sex-linked DNA markers to identify putative BAC clones, which were confirmed by Southern hybridization. We then constructed preliminary BAC contigs based on Southern hybridization of all BAC ends to marker-positive BAC clones. The ends of the terminal BACs from each contig were used to re-screen the BAC library to extend the contigs. In parallel, all BACs in the MSY region were digested with *Hin*dIII and fingerprinted by separation on agarose gels. Gel images were analysed with Image3 software (Sanger Institute) and Fingerprinting Contigs (FPC) for confirmation of the assembled contigs.

Development of *Carica papaya* sex-linked markers (cpsm)

The 225 sex co-segregating AFLP markers were surveyed to identify 85 markers that were easily resolved from neighbouring bands and were within the optimum size of 200–500 base pairs to be used as probes for physical mapping. These markers were excised from AFLP gels to develop *Carica papaya* sex-linked markers (cpsm). Sixty-three isolated cpsm fragments were reamplified with the original AFLP primers and confirmed by running the PCR products and the original AFLP products in a sequencing gel.

Test sequences in the non-recombining region

Twenty-six sequences from the non-recombining region were tested to see whether they are conserved or divergent. Nineteen test sequences were derived from the 24 cpsm markers used for physical mapping by sequencing the cpsm markers and by designing pairs of primers to amplify female, hermaphrodite and male genomic DNA. The other seven test sequences were derived from three BAC ends (cpbe45, cpbe54 and cpbe55) and from the original four RAPD-derived SCAR markers (W11, T12, PSDM and Nafp).

Subcloning and sequencing of BAC clones in the MSY

To examine the nature of DNA sequences in the MSY of the papaya primitive Y chromosome, we selected non-redundant BAC clones for sequencing. Twenty-five BACs were digested with *Hin*dIII and subcloned into the pPCR-Script Amp vector (Stratagene). Forty subclones were grown out for each BAC and their DNA was isolated. After *Hin*dIII digestion to release the inserts, 10–15 subclones of each BAC with different insert sizes were sequenced from both ends.

Gene density estimation

We randomly sequenced 628 reads, which were sampled roughly equally from 25 non-redundant BAC subclones totalling 513 kb. These were used for a blast search of the National Center for Biotechnology Information (NCBI) nucleotide sequence databases nr and est_others using BLASTN (http://www.ncbi.nlm.nih.gov/BLAST/). Functional gene sequence matches were accepted with *E*-value thresholds of less than e^{-10}. Eight hundred papaya genome survey sequences (C. Town, personal communication) totalling 605 kb were also used in a BLASTN search. We eliminated 14 mitochondrial DNA and 102 chloroplast DNA sequences, and used 684 sequences totalling 517 kb for estimating papaya genome-wide gene density.

Estimation of the sex chromosome size in papaya and *S. latifolia*

The size of the papaya sex chromosomes is estimated on the basis that the nine pairs of chromosomes have similar sizes by cytological assessment[16] and that the genome size is 372 Mb. The estimated size of the X (400 Mb) and Y (570 Mb) chromosomes in *S. latifolia* is based on the physical measurement of each chromosome from its karyotype and the female and male genome size of 5,529 Mb/2C and 5,645 Mb/2C (ref. 29), respectively. The proportion of X and Y chromosome arms in *S. latifolia* has been described[30]. There are no data on the length of the pseudoautosomal region (PAR) in *S. latifolia*, and this is a rough estimate (and thus not drawn to scale in Fig. 3).

Received 4 August; accepted 13 November 2003; doi:10.1038/nature02228.

1. Westergaard, M. The mechanism of sex determination in flowering plants. *Adv. Genet.* **9,** 217–281 (1958).
2. White, M. J. D. in *Insect Reproduction* (eds Leather, S. R. & Hardie, J.) 57–94 (Cambridge Univ. Press, Cambridge, 1973).
3. Ohno, S. *Sex Chromosome and Sex-linked Genes* (Springer, Berlin, 1967).
4. Muller, M. The relation of recombination to mutational advance. *Mutat. Res.* **1,** 2–9 (1964).
5. Rice, W. R. Genetic hitch-hiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116,** 161–167 (1987).
6. Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423,** 825–837 (2003).
7. Hofmeyr, J. D. J. Genetical studies of *Carica papaya* L. *S. Afr. Dept Agric. Sci. Bull.* **187,** 64 (1938).
8. Storey, W. B. Segregation of sex types in *Solo papaya* and their application to the selection of seed. *Am.*

Soc. Horticult. Sci. Proc. **35,** 83–85 (1938).
9. Hofmeyr, J. D. J. Some genetic and breeding aspects of *Carica papaya*. *Agronomia Tropical* **17,** 345–351 (1967).
10. Storey, W. B. in *The Evolution of Crop Plants* (ed. Simmonds, N. W.) 21–24 (Longman, London, 1976).
11. Horovitz, S. & Jiminez, H. Cruzamientos interespecificos e intergenericos en Caricaceaes y sus implicationes fitotecnias. *Agronomia Tropical* **17,** 353–359 (1967).
12. Sondur, S. N., Manshardt, R. M. & Stiles, J. I. A genetic linkage map of papaya based on randomly amplified polymorphic DNA markers. *Theor. Appl. Genet.* **93,** 547–553 (1996).
13. Ma, H. *et al.* High-density linkage mapping revealed suppression of recombination at the sex determination locus in papaya. *Genetics* (in the press).
14. Deputy, J. C. *et al.* Molecular marker for sex determination in papaya (*Carica papaya* L.). *Theor. Appl. Genet.* **106,** 107–111 (2002).
15. Ming, R. *et al.* Construction and characterization of a papaya BAC library as a foundation for molecular dissection of a tree-fruit genome. *Theor. Appl. Genet.* **102,** 892–899 (2001).
16. Storey, W. B. The botany and sex relations of the papaya. *Hawaii Agricult. Exp. Station Bull* **87,** 5–22 (1941).
17. Tanksley, S. D., Miller, J. C., Paterson, A. H. & Bernatzky, R. *Proc. 18th Stadler Genet. Symp.* 157–173 (Plenum, New York, 1988).
18. Charlesworth, B. The evolution of sex chromosomes. *Science* **251,** 1030–1033 (1991).
19. Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286,** 964–967 (1999).
20. Iwase, M. *et al.* The *amelogenin* loci span an ancient pseudoautosomal boundary in diverse mammalian species. *Proc. Natl Acad. Sci. USA* **100,** 5258–5263 (2003).
21. Charlesworth, D. Plant sex determination and sex chromosomes. *Heredity* **88,** 94–101 (2002).
22. Filatov, D. A., Moneger, F., Negrutiu, I. & Charlesworth, D. Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. *Nature* **404,** 388–390 (2000).
23. Chiu, C. T. *Study on Sex Inheritance and Horticultural Characteristics of Hermaphrodite Papaya* Thesis, Nat. Pingtung Univ. Sci. Technol., Taiwan (2000).
24. Storey, W. B. Genetics of the papaya. *J. Hered.* **44,** 70–78 (1953).
25. Okada, S. *et al.* The Y chromosome in the liverwort *Marchantia polymorpha* has accumulated unique repeat sequences harboring a male-specific gene. *Proc. Natl Acad. Sci. USA* **98,** 9454–9459 (2001).
26. Bachtrog, D. Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in *Drosophila*. *Nature Genet.* **34,** 215–219 (2003).
27. Charlesworth, B. & Charlesworth, D. A model for the evolution of dioecy and gynodioecy. *Am. Nat.* **112,** 975–997 (1978).
28. Yang, Z. N. & Mirkov, T. E. Isolation of large terminal sequences of BAC inserts based on double restriction enzyme digestion followed by anchored PCR. *Genome* **43,** 412–415 (2000).
29. Siroky, J., Lysak, M. A., Dolezel, J., Kejnovsky, E. & Vyskot, B. Heterogeneity of rDNA distribution and genome size in *Silene* spp. *Chromosome Res.* **9,** 387–393 (2001).
30. Lengerova, M., Moore, R. C., Grant, S. R. & Vyskot, B. The sex chromosomes of *Silene latifolia* revisited and revised. *Genetics* **165,** 935–938 (2003).

Supplementary Information accompanies the paper on **www.nature.com/nature**.

Acknowledgements We thank D. Charlesworth for comments on the manuscript; R. Perl-Treves for discussions; S. Ancheta, G. Asmus and L. Poland for technical assistance; R. Manshardt for providing an F_2 population for fine-mapping; and H. Albert, M. Moore, R. Osgood, B. Vyskot and S. Whalen for reviewing the manuscript. This work was supported by a United States Department of Agriculture Agricultural Research Service (USDA-ARS) Cooperative Agreement with the Hawaii Agriculture Research Center, and a subaward to R. M. and A.H.P. to produce and to characterize the BAC library.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to R.M. (rming@harc-hspa.com).
The sequences are deposited in GenBank under accession codes CG680451–CG681045 and AY428881–AY428946.

Sleep inspires insight

Ullrich Wagner[1], Steffen Gais[1], Hilde Haider[2], Rolf Verleger[3] & Jan Born[1]

[1]*Department of Neuroendocrinology, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany*
[2]*Institute of Psychology, University of Cologne, Gronewaldstrasse 2, 50931 Cologne, Germany*
[3]*Department of Neurology, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany*

Insight denotes a mental restructuring that leads to a sudden gain of explicit knowledge allowing qualitatively changed behaviour[1,2]. Anecdotal reports on scientific discovery suggest that pivotal insights can be gained through sleep[3]. Sleep consolidates recent memories[4–6] and, concomitantly, could allow insight by changing their representational structure. Here we show a facilitating role of sleep in a process of insight. Subjects performed a cognitive task requiring the learning of stimulus–response sequences, in which they improved gradually by increasing response speed across task blocks. However, they could also improve abruptly after gaining insight into a hidden abstract rule underlying all sequences. Initial training establishing a task representation was followed by 8 h of nocturnal sleep, nocturnal wakefulness, or daytime wakefulness. At subsequent retesting, more than twice as many subjects gained insight into the hidden rule after sleep as after wakefulness, regardless of time of day. Sleep did not enhance insight in the absence of initial training. A characteristic antecedent of sleep-related insight was revealed in a slowing of reaction times across sleep. We conclude that sleep, by restructuring new memory representations, facilitates extraction of explicit knowledge and insightful behaviour.

The idea that sleep can trigger the gain of insight is associated with the names of famous scientific discoverers[3]. For example, Nobel prize winner Loewi reported that he woke up with the essential idea for an experimental confirmation of his theory of chemical neurotransmission. Mendeleyev, who laid out the periodic table of chemical elements, reported that his understanding of the critical rule underlying it emerged out of a dream following unsuccessful puzzling with the symbols of the elements. Recent studies in animals and humans provided evidence for the concept that neuronal representations of task stimuli and responses acquired during wakefulness become reactivated during subsequent sleep[7–10]. This reprocessing of representations is considered to underlie the consolidating effect of sleep on memory[4,11–15], but could also be accompanied by restructuring these representations in memory to enable insight. Here, we tested whether sleep influences task representations in memory such that the gain of insight is facilitated. This was expressed in the extraction of explicit knowledge of a hidden abstract rule in stimulus–response sequences learned previously under implicit conditions.

To grasp the inherently unpredictable phenomenon of insight experimentally, we used a modified version of the Number Reduction Task (NRT; Fig. 1a) originally developed by Thurstone and Thurstone[16–18]. Based on continuous monitoring of subjects' behavioural responses, the task allows the exact determination of the time point when insight occurs, that is, when explicit knowledge of a hidden abstract rule is gained, leading to an abrupt, qualitative shift in responding. On each trial of the task, subjects were asked to transform a given string of eight digits into a new string through a stepwise digit-by-digit application of two simple rules to reach a specific digit indicating the final solution to this string. With increasing practice, the subject's responses become gradually faster on this task. Most important, however, a hidden rule was implemented in the digit strings, which was not mentioned to the subjects and was therefore initially processed at an implicit level without awareness. The time point when a subject gained insight into this rule could be determined precisely because at this time he/she would begin to cut short sequential responding to confirm the final solution in advance. All subjects were first trained on three task blocks to induce mental representations of the task that still remained implicit with regard to the hidden rule during this period. The training period was then followed by an 8-h interval of (1) nocturnal sleep, (2) nocturnal wakefulness, or (3) daytime wakefulness (Fig. 1b). Subsequently, subjects were retested on ten blocks.

Sleep more than doubled the probability of gaining insight into the hidden rule compared to wakefulness. In the sleep group, thirteen out of 22 subjects (59.1%) gained insight at retesting, compared to five subjects (22.7%) in either wake group ($\chi^2 = 8.54$, degrees of freedom (d.f.) $= 2$, $P = 0.014$; Fig. 2). For subjects gaining insight, the time point of its occurrence (number of blocks after beginning of retesting) did not differ significantly between groups (sleep, 4.5 ± 0.8 (mean $\pm$ s.e.m.); wake-night, 6.8 ± 1.5;