Nuevo Point, CA, and released offshore in Monterey Bay. After release, the seals returned to instrumentation sites where the data and videos were retrieved.
18. The experimental setup and instrumentation for the dolphin studies are described in (*25*). The instrument pack was neutrally buoyant and weighed 8 kg in air. Twenty experimental dives from 50 to 110 m were conducted in open water.
19. Blue whale studies used CRITTERCAM instrumentation (*36*) attached with a low-profile silicon suction cup (22 cm diameter). The cup released after a predetermined interval through the dissolution of a corrosible magnesium plug. The blue whale (length = 22 to 25 m) had been individually identified photographically during 1990–98 along the California coast. It was considered an adult of at least 10 years in age.
20. Gliding was defined as periods exceeding 3 to 12 s in which no locomotor movements occurred and flippers or flukes were aligned along the body axis. Deployments involving forward-facing cameras on Weddell seals also used a tail-mounted ±2-g, single-axis accelerometer (Ultramarine Instruments, Galveston, TX) to assess stroking activity. Head movements of the blue whale were considered indicative of stroke activity because of counter movements of the head and tail in swimming cetaceans (*25, 37*). Videotapes were reviewed at normal speed, except for the blue whale; cycling rate was increased sevenfold to facilitate analyses of the exceptionally slow movements of the whale.
21. R. W. Davis, T. M. Williams, G. L. Kooyman, *Physiol. Zool.* **58**, 590 (1985).
22. M. A.Castellini, G. L. Kooyman, P. J. Ponganis, *J. Exp. Biol.* **165**, 181 (1992).
23. F. E. Fish, *J. Exp. Biol.* **201**, 2867 (1998).
24. ———, S. Innes, K. Ronald, *J. Exp. Biol.* **137**, 157 (1988).
25. R. C. Skrovan, T. M. Williams, P. S. Berry, P. W. Moore, R. W. Davis, *J. Exp. Biol.* **202**, 2749 (1999).
26. P. M. Webb, D. E. Crocker, S. B. Blackwell, D. P. Costa, B. J. Le Boeuf, *J. Exp. Biol.* **201**, 2349 (1998).
27. S. H. Ridgway, B. L. Scronce, J. Kanwisher, *Science* **166**, 1651 (1969).
28. S. H. Ridgway and R. Howard, *Science* **206**, 1182 (1979).
29. G. L. Kooyman, D. D. Hammond, J. P. Scholander, *Science* **169**, 82 (1970).
30. Because seals exhale before diving, the relative contribution of lung compression to changes in buoyancy during diving is unknown for pinnipeds. The large, incompressible blubber layer of pinnipeds is less dense than seawater and may represent a significant component of the upward buoyant force (*26*).
31. D. Weihs, *J. Theor. Biol.* **48**, 215 (1974).
32. R. Stephenson, *J. Exp. Biol.* **190**, 155 (1994).
33. J. R. Lovvorn, D. A. Croll, G. A. Liggins, *J. Exp. Biol.* **202**, 1741 (1999).
34. Stroking dives, as determined from accelerometer data, were characterized by the absence of prolonged (>10 s) gliding periods and stroke-and-glide or continuous flipper movements. Gliding dives incorporated prolonged periods of gliding that ranged from 10.7 to 52.2% (mean = 36.0 ± 4.8%, n = 9; 58 to 803 s) of total dive duration.
35. R. W. Davis *et al.*, *Science* **283**, 993 (1999).
36. G. Marshall, *Mar. Technol. Soc. J.* **32**, 11 (1998).
37. J. J. Videler and P. Kamermans, *J. Exp. Biol.* **119**, 265 (1985).
38. Supported by Office of Naval Research (ONR) grant N00014-95-1-1023 and NSF Polar Programs grant OPP-9618384 to T.M.W., and ONR grant N00014-99-1-0192 to D.A.C. Blue whale research was supported through National Geographic Television. We thank G. Marshall for development of CRITTERCAM used in the blue whale study and W. Hagey for instrumentation used in the seal and dolphin studies. We also thank E. Roscow and T. Tinker for assistance with the illustrations. Animal studies were approved by individual institutional Animal Use Committees using NIH guidelines.

29 December 1999; accepted 8 March 2000

Identification of a Coordinate Regulator of Interleukins 4, 13, and 5 by Cross-Species Sequence Comparisons

G. G. Loots,[1,2] R. M. Locksley,[3] C. M. Blankespoor,[1] Z. E. Wang,[3] W. Miller,[4] E. M. Rubin,[1]* K. A. Frazer[1]*

Long-range regulatory elements are difficult to discover experimentally; however, they tend to be conserved among mammals, suggesting that cross-species sequence comparisons should identify them. To search for regulatory sequences, we examined about 1 megabase of orthologous human and mouse sequences for conserved noncoding elements with greater than or equal to 70% identity over at least 100 base pairs. Ninety noncoding sequences meeting these criteria were discovered, and the analysis of 15 of these elements found that about 70% were conserved across mammals. Characterization of the largest element in yeast artificial chromosome transgenic mice revealed it to be a coordinate regulator of three genes, *interleukin-4*, *interleukin-13*, and *interleukin-5*, spread over 120 kilobases.

Computational methods for recognizing coding sequences in genomic DNA are capable of detecting most genes; however, identifying regulatory elements in the 95% of the genome composed of noncoding sequences is currently a substantial challenge. Regulatory sequences tend to be highly conserved among mammals (*1*), suggesting comparative sequence analysis as a strategy for their identification (*2*). Extensive studies focusing on understanding the regulation of three biomedically important cytokines [*interleukin-4* (*IL-4*), *IL-13*, and *IL-5*] clustered at human 5q31 have indicated that these genes are coordinately coactivated in type 2 T helper (T_H2) cells (*3*). Transgenic mice propagating human 5q31 yeast artificial chromosomes (YACs) appropriately regulate the human *IL-4*, *IL-13*, and *IL-5* transgenes in murine T_H2 cells, independent of the site of integration (*4*). These data indicate that the elements regulating the coordinate tissue and temporal expression of these cytokines are conserved in humans and mice. A comparative sequence-based approach was employed to discover distant regulatory sequences involved in controlling the complex expression patterns of the 5q31 cytokines.

Database searches combined with GenScan predictions identified 23 putative genes (including *IL-4*, *IL-13*, and *IL-5*) in an ~1-Mb human 5q31 region and determined that the order and orientation of all but one of these genes are conserved in the murine chromosome 11 orthologous region (Fig. 1) (*5–7*). Comparative analysis of these human and mouse sequences (*8*) focused on discovering distant regulatory sequences in the region, based on the observation that these elements are typically composed of long sequences [≥100 base pairs (bp) in length] that are highly conserved among mammals (≥70% identity) (*1*). A total of 245 conserved elements fitting these criteria was identified, of which 155 overlapped with coding sequences (defined as sequences present in mature RNA transcripts) (*9*), and the remaining 90 were defined as noncoding (Figs. 1 and 2) (*10*). Of the 90 conserved noncoding elements, 46% were in introns, 9% were within 1 kb of the 5′ and 3′ ends of an identified transcript, and 45% were in intergenic regions >1 kb from any known gene. Many of the noncoding elements were found in clusters, such as in the intergenic region between *organic cation transporter 1* (*OCTN1*) and *P4–hydroxylase alpha* (*II*), suggesting that they may be working cooperatively as a functional unit (Fig. 1). One of the conserved noncoding sequences (CNSs), CNS-7, located in the intergenic region between *granulocyte-macrophage colony-stimulating factor* (*GM-CSF*) and *IL-3*, had previously been identified experimentally as an enhancer controlling the coregulation of these two cytokines (*11*). This finding supports the choice of the criteria used in this sequence-based approach to iden-

[1]Genome Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [2]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA. [3]Departments of Medicine and Microbiology/Immunology and Howard Hughes Medical Institute, University of California, San Francisco, CA 94143, USA. [4]Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA 16802, USA.

*To whom correspondence should be addressed. E-mail: kafrazer@lbl.gov (K.A.F.) and emrubin@lbl.gov (E.M.R.)

tify biologically relevant noncoding sequences.

Fifteen of the CNSs were assessed for their presence in other vertebrates and their copy number in the human genome. Degenerate primer pairs of the elements were used to amplify genomic DNA of other vertebrates, and the resulting products were sequenced (*12*). Ten of the elements were highly conserved in at least two mammals in addition to humans and mice (Fig. 1 and Table 1). Twelve elements appeared unique in the human genome, as determined by low-stringency Southern blot hybridizations (*13*). Of the noncoding sequences examined, ~70% are conserved across mammals and unique in the human genome, features commonly noted in experimentally identified distant regulatory elements (*1*).

The largest conserved noncoding sequence, CNS-1 (401 bp), which is located in the intergenic region (~13 kb) between *IL-4* and *IL-13*, was chosen for in-depth functional analysis on the basis of several features that suggest it might be a distant regulatory element. CNS-1 demonstrates a high degree of conservation across mammals (~80% identity in mice, humans, cows, dogs, and rabbits) (Table 1), contrasting sharply with the relatively low conservation observed in the coding regions of the flanking genes, *IL-4* and *IL-13* (~50% identity between humans and mice) (Fig. 2). This element is single copy in the human genome and has been conserved during evolution not only with regards to sequence but also to genomic location [mapped in dogs, baboons, humans, and mice to the *IL-4* through *IL-13* intergenic region (*14*)]. Binding sites for transcription factors known to regulate the expression of *IL-4* and *IL-13* were not found in CNS-1 (*15*); however, we determined that CNS-1 overlaps with two (T_H2 cell specific) of the eight deoxyribonuclease (DNase) I hypersensitive sites (HSs) previously localized in the *IL-4* through *IL-13* region (*16*) (Fig. 2).

The biological properties of CNS-1 were characterized through the creation and analysis of multiple lines of mice bearing a 450-kb human YAC transgene (Fig. 1) either containing or lacking the CNS-1 element. To reduce the uncertainties of comparing different founder lines of transgenic mice, we inserted *loxP* sites (*17*) into the YAC transgene flanking the CNS-1 element and introduced this modified YAC into the genome of mice (three separate founder lines were created). To delete the CNS-1 element, we bred the human YAC transgenic mice from each of the founder lines with mice expressing a Cre recombinase transgene (*18*). This resulted in the generation of two lines of transgenic mice derived from each founder line, one with a CNS-1–containing YAC transgene ($CNS\text{-}1^{wt}$) and one with a YAC transgene in which the CNS-1 element had been deleted ($CNS\text{-}1^{del}$). The deletion of CNS-1 was the only rearrangement in the human *IL-4* through *IL-13* region detected by Southern blot analysis (*19*). YAC copy number was the same in the paired $CNS\text{-}1^{wt}$ and $CNS\text{-}1^{del}$ transgenic animals as determined by both Southern blot analysis (*19*) and fluorescent in situ hybridization (FISH) (Fig. 3) (*20*). These findings suggest that the paired $CNS\text{-}1^{wt}$ and $CNS\text{-}1^{del}$ transgenic mice are genetically identical in every aspect except for the presence or absence of the CNS-1 element.

To assess whether CNS-1 affects expression of the human 5q31 cytokines, we isolated highly purified naïve $CD4^+$ T cells from the spleen and lymph nodes of paired $CNS\text{-}1^{wt}$ and $CNS\text{-}1^{del}$ transgenic mice and stimulated the cells in vitro under conditions that favor development of either T_H1 or T_H2 cells (*21, 22*). Neither the *IL-4* nor *IL-13* human transgenes were expressed during T_H1 differentiation in either the $CNS\text{-}1^{wt}$ or $CNS\text{-}1^{del}$ transgenic cells, suggesting that CNS-1 was not required for the repression of these cytokines in T_H1 cells. When T cells were stimulated to promote T_H2 differentiation, the $CNS\text{-}1^{del}$ transgenic cells developed less than half as many human IL-4–producing cells and less than a third as many human IL-13–producing cells as the $CNS\text{-}1^{wt}$ transgenic mice did (Fig. 4, A and B) (*22*). In the T_H2 cells that expressed human IL-4 and IL-13,

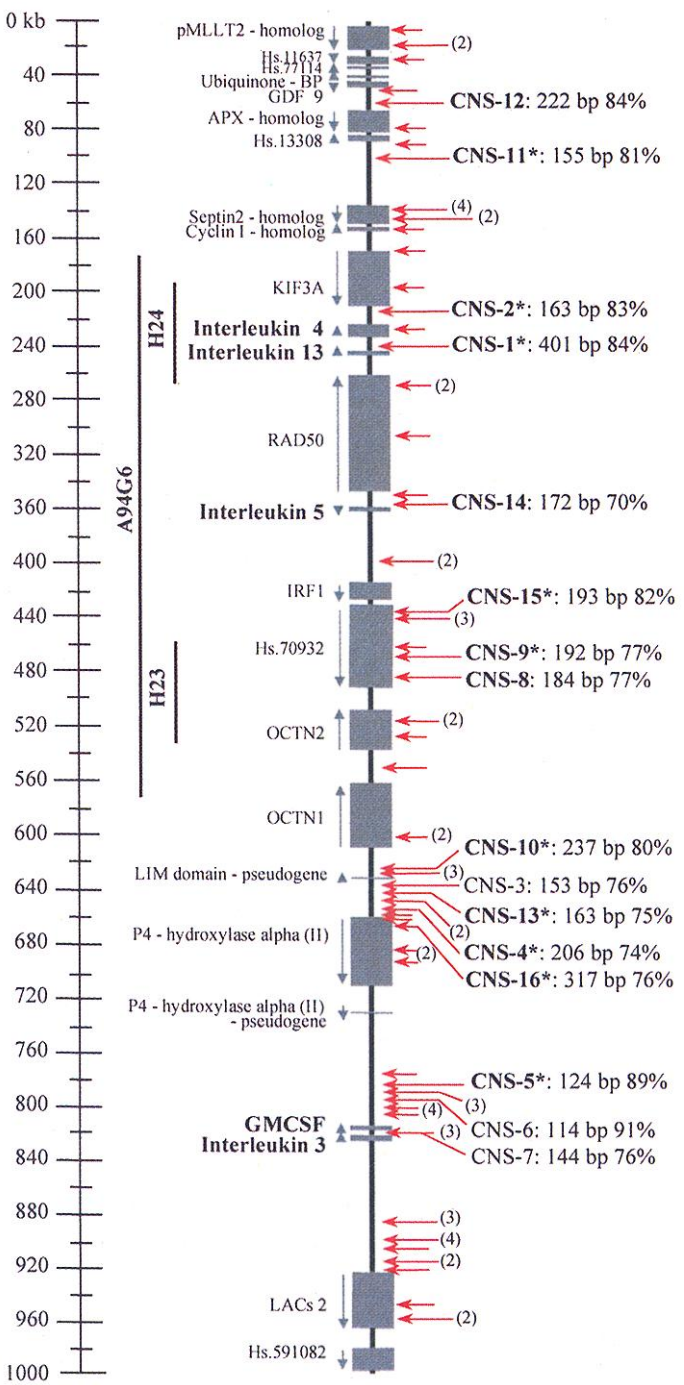

Fig. 1. Physical map of the 1-Mb human 5q31 region. The locations of the 23 genes in the interval (*7*) (gray boxes), their direction of transcription (gray vertical arrows), and the locations of all CNSs (red horizontal arrows) are shown. A number next to a horizontal arrow indicates that more than one CNS was found at that location. The lengths and percent identities of the human and mouse alignments of the 15 noncoding sequences investigated in this study are given. The 12 elements determined to be single copy in the human genome are in bold. Noncoding sequences conserved in other vertebrates (Table 1) are marked with an asterisk. The locations of the human YAC (A94G6) and P1 clones (H23 and H24) (vertical bars) used in this study are shown.

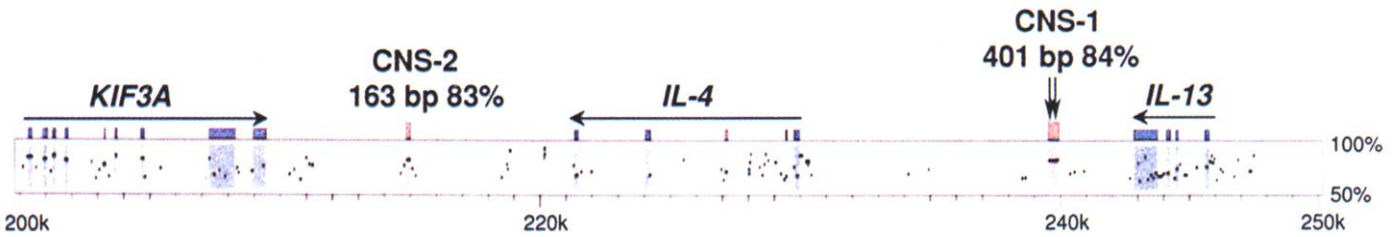

Fig. 2. Percent identity plot displaying locations for CNS elements in the *IL-4* through *IL-13* region. Conserved sequences are plotted in relation to their coordinates in the human genome (horizontal axes), and their percent identities are indicated on the vertical axis (*9*). Sequences defined as coding (blue) (*8*) and noncoding (≥100 bp and ≥70%) (red) are indicated. A horizontal arrow indicates the direction of transcription for each gene. The two DNase I HSs that overlap with CNS-1 are shown as vertical arrows. k, kilobases.

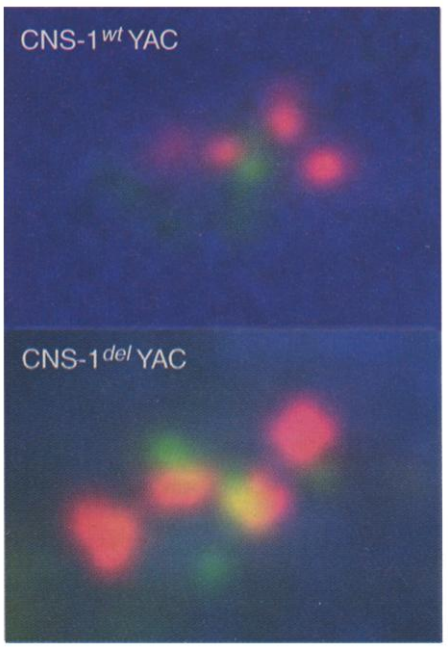

Fig. 3. YAC copy number determined by FISH. Interphase nuclei of spleen cells isolated from paired (**top**) CNS-1^{wt} and (**bottom**) CNS-1^{del} YAC transgenic mice (line 1) hybridized with two P1 probes [H23 (green) and H24 (red)] (Fig. 1) as described (*19*) are shown. Similar hybridization patterns were obtained for line 3.

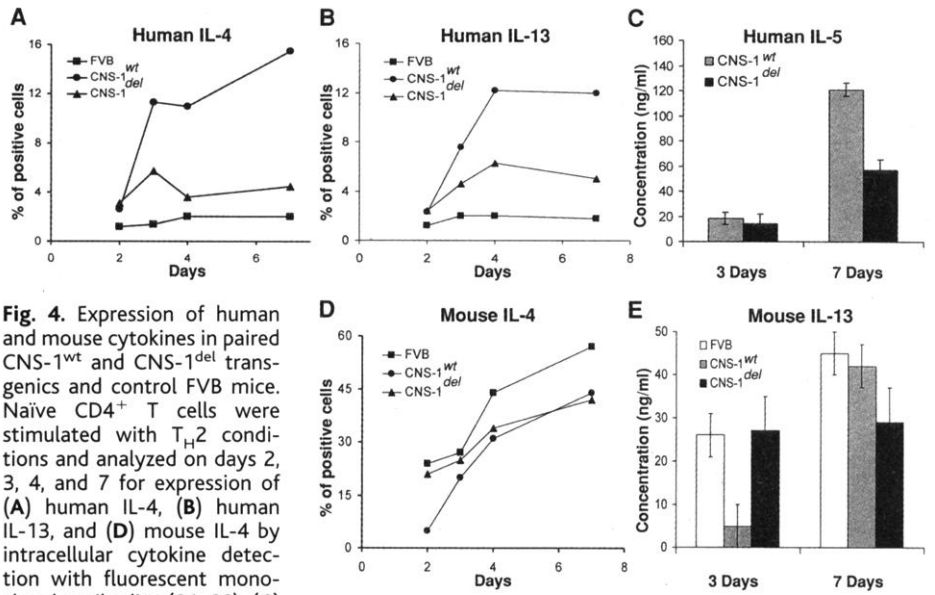

Fig. 4. Expression of human and mouse cytokines in paired CNS-1^{wt} and CNS-1^{del} transgenics and control FVB mice. Naïve CD4$^+$ T cells were stimulated with T_H2 conditions and analyzed on days 2, 3, 4, and 7 for expression of (**A**) human IL-4, (**B**) human IL-13, and (**D**) mouse IL-4 by intracellular cytokine detection with fluorescent monoclonal antibodies (*21*, *22*). (**C**) Human IL-5 and (**E**) mouse IL-13 protein levels were determined by ELISA with supernatants of activated T cells collected at either 3 or 7 days (following restimulation) after stimulation of naïve cells (*21*, *22*). Bars represent means and standard errors of the means.

however, the amounts produced per cell, assessed by the mean fluorescence intensity, were the same in the paired human CNS-1^{wt} and CNS-1^{del} YAC transgenics. The observation that CNS-1 influences the number of T_H2 cells expressing *IL-4* and *IL-13* but does not influence their levels of expression per cell suggests that this element does not act as a classical enhancer but rather appears to be involved in modulating chromatin structure. Production of human IL-5 was also significantly reduced in the CNS-1^{del} transgenic cells as compared with the paired CNS-1^{wt} transgenic cells (Fig. 4C). The lack of suitable antibodies for flow cytometry detection of human IL-5 precluded us from distinguishing whether the reduced expression of human IL-5 was due to a decrease in the number of cells producing it or due to a decrease in the amounts produced by individual T_H2 cells. All three paired CNS-1^{wt} and CNS-1^{del} YAC transgenic lines were examined, and the absence of CNS-1 consistently resulted in a decrease in the number of T_H2 cells expressing human IL-4 and IL-13.

Transcript levels of human *IL-4*, *IL-13*, and *IL-5*, as well as the two noninterleukin genes closest to CNS-1, *kinesin family member 3A* (*KIF3A*) and *RAD50*, were quantified in the paired CNS-1^{wt} and CNS-1^{del} transgenic mice (*23*). The mRNA levels of the three human cytokines were reduced in CNS-1^{del} T_H2 cells, reflecting the differences observed in the protein levels and indicating that CNS-1 acts through its effect on the transcriptional activity of these genes. Expression levels of human *KIF3A* and *RAD50* were essentially the same in the brains, hearts, kidneys, livers, and isolated T_H2 cells of the paired CNS-1^{wt} and CNS-1^{del} transgenic mice. *RAD50* is a large gene (spanning 87 kb) located between CNS-1 and *IL-5* (Fig. 1). The fact that the CNS-1^{del} transgenic mice produce substantially less human *IL-5*, but unaltered amounts of *RAD50*, suggests that CNS-1 acts over a large genomic interval to specifically affect the expression of T_H2 cell–specific cytokines.

The initial production of murine IL-4 and IL-13 was significantly less in CNS-1^{wt} transgenic T_H2 cells, whereas at later time points, CNS-1^{wt} and CNS-1^{del} transgenic T_H2 cells expressed comparable amounts of these cytokines (Fig. 4, D and E). Human IL-4 and IL-13 do not modulate murine T_H2 cell development in vitro, as determined with CD4$^+$ T cells from nontransgenic mice (*24*). These data indicate that the initial reduction in murine IL-4 and IL-13 production in the CNS-1^{wt} transgenic cells is probably not due to the expression of the human IL-4 and IL-13 transgenes, suggesting that a competitive interaction may exist between the human CNS-1 element and a trans-acting murine factor(s). The observation that, at subsequent time points, the murine cytokines are expressed at equivalent levels in CNS-1^{wt} and CNS-1^{del} transgenic cells may be due either to alterations in the levels of such putative trans factor(s) with increasing cell division or to preferential expansion or survival of murine IL-4– and IL-13–producing cells.

Our study illustrates the utility of comparative sequence analysis in the identification of

Table 1. Cross-species sequence analysis of 15 conserved noncoding elements. Human and mouse sequence alignments of each element were inspected, and primer pairs were chosen from the most conserved regions. Hence, for each CNS, the size of the PCR amplified product(s) is smaller than the size of the element indicated in Fig. 1. Numbers in parentheses are percent identity. Dashes indicate that PCR did not amplify a product.

Conserved noncoding sequence	Product base-pair size								
	Human	Cow	Dog	Pig	Rabbit	Rat	Mouse	Chicken	Fugu
CNS-1	224	220 (86)	218 (80)	229*	200 (72)	214 (75)	214 (76)	266*	155*
CNS-2	104	104 (90)	105 (93)	104 (89)	101 (90)	105 (85)	104 (85)	192*	245*
CNS-3	131	–	–	–	–	–	131 (73)	–	–
CNS-4	77	77 (88)	–	–	–	77 (87)	77 (87)	–	–
CNS-5	157	–	157 (82)	–	–	159 (84)	159 (81)	–	–
CNS-6	148	–	–	–	–	–	139 (87)	–	–
CNS-8	122	–	–	–	–	–	122 (79)	–	–
CNS-9	124	125 (88)	125 (91)	125 (88)	125 (92)	124 (86)	124 (86)	120*	111*
CNS-10	213	214 (82)	–	210 (84)	209 (83)	211 (79)	211 (77)	–	–
CNS-11	120	–	119 (89)	119 (90)	120 (83)	119 (80)	119 (79)	–	–
CNS-12	200	–	–	–	–	194†	194 (78)	–	–
CNS-13	314	315 (73)	295*	317 (75)	–	324 (67)	327 (68)	–	473*
CNS-14	128	–	–	–	–	–	128 (60)	–	–
CNS-15	127	125 (90)	125 (91)	125 (90)	–	127 (86)	127 (86)	–	88*
CNS-16	109	109 (74)	–	–	109 (84)	111 (83)	109 (80)	–	–

*Only primer sequences were conserved. †Amplified PCR product was not sequenced.

distant regulatory elements. Of the 15 human and mouse conserved noncoding elements we examined, most are also present in other mammals. These data, in combination with analysis of the orthologous human 5q31 interval in dogs revealing a strikingly similar pattern of CNS elements (*25*), suggest that most of the human and mouse CNS elements we identified have been actively conserved because of a biological function. CNS-1 appears to be involved in gene activation by modulating chromatin structure, a function for which there are no standard in vitro assays, suggesting that this element would have been difficult, if not impossible, to identify with traditional experimental methods. The use of a 450-kb YAC transgene containing eight human genes, each potentially serving as "a reporter," allowed us to determine that CNS-1 regulates some, but not all, of the genes in a large genomic interval (120 kb) on human 5q31. These regulatory features of CNS-1 reveal the complexity of long-range regulatory elements and the power of comparative biology in discovering and deciphering the properties of such elements.

References and Notes

1. Q. Li, S. Harju, K. R. Peterson, *Trends Genet.* **15**, 403 (1999).
2. R. C. Hardison, J. Oeltjen, W. Miller, *Genome Res.* **7**, 959 (1997); W. Miller, PipMaker (additional program information is available at http://globin.cse.psu.edu/pipmaker).
3. A. Abbas, K. M. Murphy, A. Sher, *Nature* **383**, 787 (1996).
4. D. J. Symula *et al.*, *Nature Genet.* **23**, 241 (1999).
5. Human 5q31 sequence (clones: H14, H13, H23, H11, H16, H21, H18, H15, H24, H17, H26, H25, H81, H22, H20, and H82) was obtained (for a list of human clones submitted to GenBank, see http://www-hgc.lbl.gov/human-p1s.html) and assembled into nine contigs of the following sizes (given in the directionality of centromere to telomere): 380.0, 13.8, 155.3, 106.9, 155.5, 5.4, 8.9, 32.3, and 90.0 kb. On the basis of the physical map of the region, the eight gaps have estimated sizes of 12, 4, 5, 3, 16, 2, 5, and 5 kb, respectively. Seven mouse chromosome 11 bacterial artificial chromosomes (BACs) were isolated from California Institute of Technology library CitbCJ7 and sequenced in either draft or finished format: 24OC4 (202 kb), 3.5× draft; 111I81 (151 kb), AC005742 finished; 219O10 (204 kb), 5.0× draft; 32I19 (131 kb), 4.0× draft; 33O24 (150 kb), 3.5× draft; 54K15 (145 kb), 5.0× draft; and 327O23 (185 kb), AF248716 4.0× draft.
6. K. A. Frazer *et al.*, *Genome Res.* **7**, 495 (1997).
7. Human repetitive elements were masked with RepeatMasker (A. F. A. Smit and P. Green; additional program information is available at http://www.genome.washington.edu/uwgc/analysistools/repeatmask.htm). Using gapped BLAST searches (default parameters), we compared masked 5q31 sequences with sequences in the National Center for Biotechnology Information (NCBI) databases, GenPept (release 113) and UniGene (30 September 1999), and analyzed the sequences for potential coding regions with GenScan [C. Burge and S. Karlin, *J. Mol. Biol.* **268**, 78 (1997)]. The following 14 genes in the interval were previously known and localized on the basis of exact GenPept database matches (NCBI numbers are given): *Ubiquinone-binding protein* (D50369), *GDF-9* (NM_005260), *KIF3A* (AF041853), *IL-4* (A00076), *IL-13* (U10307), *RAD50* (U63139), *IL-5* (NM_000879), *IRF1* (P10914), *OCTN2* (NM_003060), *OCTN1* (NM_003059), *P4–hydroxylase alpha (II)* (NM_004199), *GM-CSF* (P04141), *IL-3* (P08700), and *Long chain fatty acyl CoA synthetase 2/KIAA0837* (AF099740/AB020644). Four genes were identified on the basis of GenScan predictions, exact expressed sequence tag (EST) matches, and their similarities to known proteins, and they are referred to as a homolog of the gene to which they had the highest BLASTX score: *pMLLT2-homolog* (NM_005935), *APX-homolog* (Q01613), *Septin2-homolog* (D86957), and *Cyclin I–homolog* (NM_006835). Of the five genes referred to by their UniGene number, one was identified solely on the basis of an exact UniGene match (Hs.70932), and four were identified by exact UniGene matches in conjunction with GenScan predictions and mouse and human conserved sequences (Hs.11637, Hs.77114, Hs.13308, and Hs.591082). Two database matches were determined to be pseudogenes: the *Lim-domain* match (X93510) was exact but only to the first 98 out of 1130 nucleotides in the mRNA, and the *P4–hydroxylase alpha (II)* pseudogene had neither an open reading frame nor an exact matching EST.
8. Alignments between the human and mouse genomic sequences were computed with a dynamic programming method, scoring each nucleotide match as 1, each mismatch as −1, and each gap of length k as -6 to $0.2k$. The percent identity plot displays the human positions and the percent identity of each segment of the alignment between successive gaps that has a length of at least 40 bp and at least 60% identity.
9. Sequences present in mature mRNAs were computationally identified either by exact GenPept matches or exact UniGene matches combined with partial GenPept matches and/or GenScan predictions of probability scores ≥0.15. Database matches include the 5′ and 3′ untranslated regions of mRNAs; therefore, some of the 155 conserved coding sequences, as defined in this study, are not translated.
10. To prevent the inclusion of RNA and RNA pseudogenes in this set of conserved noncoding sequences, we masked all tRNA and most of the known small nuclear RNA genes in the human 5q31 sequence with RepeatMasker. However, because computational screens are biased against RNA genes, it is possible that a fraction of these 90 conserved noncoding elements may actually be transcribed.
11. C. S. Osborne, M. A. Vadas, P. N. Cockerill, *J. Immunol.* **155**, 226 (1995).
12. Genomic DNA was purchased from Clontech Laboratories (Palo Alto, CA) (catalog numbers dog 6950-1, rabbit 6960-1, rat 6750-1, mouse 6650-1, human 6550-1, porcine 6651-1, bovine 6850-1, and chicken 6852-1). *Drosophila melanogaster* and *Fugu rubripes* genomic DNA were isolated with standard methods. Polymerase chain reaction (PCR) amplifications were performed as follows: 100 ng of genomic DNA from each species was mixed with 200 μM of each deoxyribonucleoside triphosphate, 1 μM of each oligonucleotide primer (5 μM for degenerate primers), 5 μl of 10× PCR buffer (PerkinElmer), and five units of AmpliTaq DNA Polymerase (PerkinElmer) in a 50-μl volume. The samples were amplified under standard PCR reaction conditions in an automated thermal cycler (PerkinElmer 9700) for a total of 35 cycles with the following primer pairs: CNS-1 forward, 5′-TGATTTCTCGGCAGCCAGGGAGGGCC-3′; CNS-1 reverse, 5′-GGTGCCTGCGTCACCTCTGACCACAC-3′; CNS-2 forward, 5′-CCTCTCAGCATTTATCTTGGGC-3′; CNS-2 reverse, 5′-AGAGCCATAANNGTGTTT-GGG-3′; CNS-3 forward, 5′-CNAGTNGNTCAGGGC-NNGATGCCCAGG-3′; CNS-3 reverse, 5′-AAGGGNG-TCTGNTCNTNCTGGAGCCTGCC-3′; CNS-4 forward, 5′-GCATGAAGNATTGNTGGCCC-3′; CNS-4 reverse, 5′-CTCTCTGGCNCTGGAACACC3′; CNS-5 forward, 5′-ACNGTTTTTNGTGTGCAGCACT-3′; CNS-5 reverse, 5′-ATTCTTTNAAAACCCCATATC-3′; CNS-6 forward, 5′-TAGNANAGTGAGGATGTCTG-3′; CNS-6 reverse, 5′-AAACCCCAGCNCTGGGCAAACAG-3′; CNS-8 forward, 5′-AAGTAAACNCTGNAAAANNTG-3′; CNS-8 reverse, 5′-CNCNNAAGTATACTTTGGAATCC-3′; CNS-9 forward, 5′-TNACTCNCAGTGACTGATNTTTG-3′; CNS-9 reverse, 5′-ATCNCCTCCNNGTNTCTTTGCAAC-3′; CNS-10 forward, 5′-CANGATGACTCAGCCAGCACAAG-3′; CNS-10 reverse, 5′-CCTNNTCTAGGAAATGGGCT-TGC-3′; CNS-11 forward, 5′-GGCAAANTGTCA-

CAATGTTC-3′; CNS-11 reverse, 5′-CTGTCANAGCCACACAGAAG3′; CNS-12 forward, 5′-TCCACATTTTCTTNCCTTTG-3′; CNS-12 reverse, 5′-GTNTCNCTGCCCTTTGATG-3′; CNS-13 forward, 5′-GGNTGAGATNCTGGAGGCTC-3′; CNS-13 reverse, 5′-GAGCAGGTCTGACNNGGGTG-3′; CNS-14 forward, 5′-TTGGCAATTCCCCTGAAA C-3′; CNS-14 reverse, 5′-AAGCTKAGYTCTGGCAGG-3′; CNS-15 forward, 5′-AAGNNTGTT GCTANGGTCACTGTG-3′; CNS-15 reverse, 5′-GCAGNTGTGGTTTTGAGANGTTCA T-3′; CNS-16 forward, 5′-CTCCCACATCCTTGGGAGGG-3′; and CNS-16 reverse, 5′-CCAGNAGCCAGGGACACACC-3′. Amplified PCR products were analyzed by gel electrophoresis on 3% Nusieve GTG agarose gels (BioWhittaker Molecular Applications, Rockland, ME), extracted with QIAquick Gel Extraction Kits (QIAGEN, Valencia, CA), and sequenced with Big Dye chemistry (PE Applied Biosystems, Foster City, CA).

13. Human genomic DNA (8 μg) (catalog number 6550-1 from Clontech Laboratories) was digested (with Pst I, Msp I, Bgl II, Pst I, and Taq$^{\alpha}$I), separated in an 0.8% agarose gel, transferred onto membrane (Hybond N^{+}, Amersham), hybridized at 52°C overnight in Church Buffer [0.5 M $NaHPO_4$ (pH 7.2), 1 mM EDTA, 7% SDS, and 1% bovine serum albumin] with the most stringent wash at 58°C [0.04 M $NaHPO_4$ (pH 7.2), 1 mM EDTA, and 1% SDS]. Probes were generated by PCR amplification of human genomic DNA using primer pairs of each CNS element and then gel purified (*12*). A CNS element was defined as single copy if none of the five lanes containing DNA (one of each restriction digest) had more than one or two bands (assumed to be a polymorphism).
14. Gridded high-density dog and baboon libraries (BACPAC Resources, Roswell Park Cancer Institute, Buffalo, NY) were screened by hybridization with CNS-1 probes generated by PCR amplification of dog and human genomic DNA (*12*), respectively. Content mapping of 9 dog BACs and 14 baboon BACs for the presence of *KIF3A*, CNS-2, *IL-4*, CNS-1, *IL-13*, and *RAD50* by PCR defined the location of CNS-1 to the *IL-4* through *IL-13* region.
15. CNS-1 sequences (human, mouse, dog, rat, cow, and rabbit) were searched (the Transcription Factor Database is available at http://transfac.gbf-braunschweig.de/TRANSFAC/index.html) for consensus binding sites of four proteins [NF-AT (nuclear factor of activated T cells), c-maf, GATA-3, and STAT6 (signal transducer and activator of transcription 6)] known to regulate *IL-4* transcription. No conserved consensus binding sites were found.
16. N. Takemoto *et al.*, *Int. Immunol.* **10**, 1981 (1998); S. Agarwal and A. Rao, *Immunity* **9**, 765 (1998).
17. YAC A94G6 (450 kb) (*6*) was retrofitted with pLys2neo vector (gift from K. Peterson) as described [B. C. Lewis, N. P. Shah, B. S. Braun, C. T. Denny, *Genet. Anal. Tech. Appl.* **9**, 86 (1992)]. The yeast shuttling vector, pRS406.CNS-1.*loxP*, was constructed as follows: A 2.4-kb Sac I fragment containing human CNS-1 was cloned into pRS406 (Stratagene) to generate pRS406.CNS-1. Oligonucleotides, LoxP–Pml I (forward 5′-GTGTAACTTCGTATAGCATACATTATACGAAGTTATCAC-3′; reverse 5′-GTGATAACTTCGTATAATGTATGCTATACGAAGTTACAC) and LoxP–Sph I (forward 5′-CTAACTTCGTATAGCATACATTATACGAAGTTATGCATG-3′; reverse 5′-CATAACTTCGTATAATGTATGCTATACGAAGTTAGCATG-3′), containing *LoxP* sequences with sticky ends were synthetically synthesized, annealed in vitro, and subcloned into Pml I and Sph I sites of the pRS406.CNS-1 vector, creating pRS406.CNS-1.*loxP*. This vector was linearized at the PfIM I site, and the pop-in/pop-out method [K. Duff, A. McGuigan, C. Huxley, F. Schultz, J. Hardy, *Gene Ther.* **1**, 1 (1993)] was used to modify the retrofitted A94G6 YAC. YAC DNA was isolated at a final concentration of ~1 ng/ml and microinjected into fertilized FVB mouse eggs using standard procedures as previously described (*6*).
18. K. Wagner *et al.*, *Nucleic Acids Res.* **25**, 4323 (1997).
19. G. G. Loots, unpublished observations.
20. For FISH, slides of lymphocytes isolated from 4- to 6-week-old F1 transgenics were prepared and hybridized with two human P1's (H23 and H24) (*5*) as described [E. D. Green *et al.*, in *Mapping Genomes*, vol. 4 of *Genome Analysis: A Laboratory Manual Series*, B. Birren *et al.*, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1999), pp. 303–413].
21. Naïve CD4^{+} T cells were sorted to >99% purity from spleen and lymph nodes of transgenics (6 to 8 weeks old) on the basis of small forward- and side-scattering characteristics and CD4^{+}, CD62L^{hi} phenotype. Cells were activated using irradiated antigen-presenting cells with monoclonal antibodies against β T cell receptor and CD28 with IL-2; for Th1 conditions, recombinant murine IL-12 and antibody to IL-4 were used, and for Th2 conditions, recombinant murine IL-4 was used.
22. D. J. Fowell *et al.*, *Immunity* **11**, 399 (1999). At periods from 2 to 7 days, CD4^{+} T cells were analyzed with flow cytometry (positive cells were gated as compared to isotype antibody controls and are based on at least 10,000 flow cytometric events) for expression of the designated murine and human cytokines by intracellular cytokine detection after 4 hours of restimulation with phorbol myristate acetate and ionomycin (PMA/IONO) (*4*). Supernatants from activated T cells were collected (either 72 hours after the primary stimulation with antibodies and irradiated antigen presenting cells or 24 hours after restimulation of 7-day-old cultures with PMA/IONO) and analyzed with enzyme-linked immunosorbent assay (ELISA) for human IL-5 and murine IL-13.
23. Total RNA was isolated with RNA-STAT-60 (TEL-TEST"B"). Five micrograms of RNA was reverse-transcribed into cDNA (Superscript *II*, Gibco), and expression levels were measured with TaqMan Syber-Green quantitative PCR assay (PerkinElmer). The sequences of the TaqMan primer pairs used to quantify mRNA are as follows: human IL-4 forward, 5′-ACAGCCTCACAGAGCAGAAGACT-3′; human IL-4 reverse, 5′-GTGTTCTTGGAGGCAGCAAAG-3′; human IL-5 forward, 5′-ATAAAAATCACCAACTGTGCACTGAA-3′; human IL-5 reverse, 5′-CAAGTTTTTGAATAGTCTTTCCACAGTAC-3′; human IL-13 forward, 5′-CAGAAGCTCCGCTCTGCAAT-3′; human IL-13 reverse, 5′-ACACGTTGATCAGGGATTCCA-3′; human KIF3A forward, 5′-CGCAGTCTCGAGAGCGTCAA-3′; human KIF3A reverse, 5′-ACACCGGGTGCGCAGA-3′; human RAD50 forward, 5′-TGTTGGCTGGCAGGATCTTT-3′; human RAD50 reverse, 5′-CGTGAGACCCGCGAATCT-3′; mouse glyceraldehyde phosphate dehydrogenase (mGAPDH) forward, 5′-GGCAAATTCAACGGCACAGT-3′; and mGAPDH reverse, 5′-CCTCACCCCATTTGATGTTAGTG-3′.
24. R. M. Locksley and Z. E. Wang, unpublished observations.
25. K. A. Frazer, unpublished observation.
26. We thank A. Nyugen, W. Dean, and K. Lewis for DNA sequencing; J. F. Cheng for isolating mouse chromosome 11 BACs; K. Frankel for assistance with sequence assembly; K. U. Wagner for providing the Cre recombinase transgenic mice; I. Plajzer-Frick for assistance with FISH; C. McArthur for flow cytometry; and K. Peterson and M. Dunaway for critical reading of the manuscript. Supported by the following grants: U.S. Department of Energy contract DE-AC0376SF00098, NIH GM-5748202 (K.A.F.), Howard Hughes Medical Institute, AI30663 and NIH HL56385 (R.M.L.), and National Library of Medicine LM05110 (W.M.).

12 November 1999; accepted 22 February 2000

Specialized Fatty Acid Synthesis in African Trypanosomes: Myristate for GPI Anchors

Yasu S. Morita, Kimberly S. Paul, Paul T. Englund*

African trypanosomes, the cause of sleeping sickness, need massive amounts of myristate to remodel glycosyl phosphatidylinositol (GPI) anchors on their surface glycoproteins. However, it has been believed that the parasite is unable to synthesize any fatty acids, and myristate is not abundant in the hosts bloodstreams. Thus, it has been unclear how trypanosomes meet their myristate requirement. Here we found that they could indeed synthesize fatty acids. The synthetic pathway was unique in that the major product, myristate, was preferentially incorporated into GPIs and not into other lipids. The antibiotic thiolactomycin inhibited myristate synthesis and killed the parasite, making this pathway a potential chemotherapeutic target.

Trypanosoma brucei causes human sleeping sickness and livestock disease in Africa. The bloodstream form of this parasite covers its surface with ~10^7 identical molecules of a glycosyl phosphatidylinositol (GPI)–anchored variant surface glycoprotein (VSG). The VSG GPI contains the fatty acid myristate (14:0) as its lipid moiety (*1*), and fatty acid remodeling reactions incorporate these myristates into the GPI precursor (*2*). Despite the large requirement for myristate, this fatty acid is not abundant in host bloodstreams (*3*), and the trypanosome was believed to be unable to either synthesize fatty acids de novo or to shorten longer fatty acids by β oxidation (*4*). This dilemma (*5*) provided a rationale for our investigation of fatty acid metabolism in trypanosomes. In studies on the cultured bloodstream form of *T. brucei* (*6*), we found that the elongation of laurate (12:0) to myristate was highly efficient, whereas the elongation of myristate to palmitate (16:0) was inefficient. We wondered why trypanosomes have a mechanism for the elongation of laurate, as there is little laurate in host bloodstreams (*3*) or in the parasite (*7*). We considered the possibility that trypanosomes can actually synthesize fatty acids de novo and that the elongation of laurate is a step in this pathway.

To test this hypothesis, we established a trypanosome cell-free fatty acid synthesis sys-

Department of Biological Chemistry, Johns Hopkins Medical School, Baltimore, MD 21205, USA.

*To whom correspondence should be addressed. E-mail: penglund@jhmi.edu