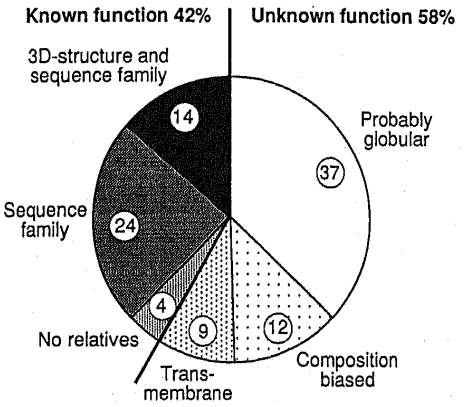
What's in a genome?

SIR — We have taken up the challenge to elucidate the function of the 182 predicted protein products derived from the complete DNA sequence of yeast chromosome III, a result of the European yeast genome project (S. Oliver *et al. Nature* **357**, 38–46; 1992). These authors report that some functional information is available for about 57 of these proteins, either determined by experiment or deduced by similarity searches in sequence databases.

We have identified the probable function of 17 additional protein products, using a combination of low-stringency sequence database searches, various ways of assessing significance, multiple sequence alignment, pattern searches and incorporation of prior knowledge about protein and domain families. The most interesting of these include a DNA polymerase of type X previously found only in mammalia, a new regulatory domain common to eukaryotes and prokaryotes (PILB), a methyltransferase, an acetolactate synthase and a GAL4-type transcriptional activator (see table). In addition, we find that 25 of the chromosome III proteins have homologues of known three-dimensional structure. Taken together, as many as 42% of all proteins of yeast chromosome III have a known or probable function and 13% have an indirectly known three-dimensional structure. Of the remaining 58%, about one-third have putative transmembrane segments (see figure).



Yeast chromosome III proteins. Information accumulated to date by all methods, experimental and theoretical. Information content increases counterclockwise. The principal diversion is between known and unknown biological function. Numbers in per cent. Composition bias indicates unusual amino-acid composition untypical of globular proteins, for example, in coiled coils. The categories are approximate, but give an impression of the current state of the art.

Extrapolating from chromosome III to the entire yeast genome, we can expect that the white, uncharted areas cover only about half of the protein function map but as much as six-sevenths of the protein structure map. As genome projects provide more and more raw sequence data, informatics methods such as those used here can cover much ground, but efficient experimental methods are needed to determine the structure and function of proteins without similarity to known families.

Peer Bork, Christos Ouzounis, Chris Sander, Michael Scharf, Reinhard Schneider & Erik Sonnhammer
European Molecular Biology Laboratory, D–6990 Heidelberg, Germany

SIMILARITY OF SELECTED CHROMOSOME III PRODUCTS TO OTHER PROTEINS

ORF	Length	Family	Closest	%id/len
YCL9c	309	Prokaryotic acetolactate synthases, small subunit	ILVH_ECOLI	36/170
YCL19w	1,347	Transposon B gene family, related to pol genes	COPI_DROME	18/505
			POLX_TOBAC	20/393
YCL20w	438	Transposon A gene family	YTY1_YEAST	49/439
YCL33c	168	Repressor of pilin promoter	PILB_NEIGO	33/110
YCL75w	146	Pol-like protein	S00954(P)	40/74
YCR14c	582	Type X DNA polymerases	DPOB_RAT	26/393
YCR23c	611	Tetracycline resistance proteins	TCR1_ECOLI	28/150
YVR26c	743	Mammalian PC1 plasma cell membrane protein phosphodiesterase family	PC1_HUMAN	38/129
			PPD1_BOVIN	38/66
YCR32w	2,167	Hypothetical protein related to C-terminal 'CDC4'-like human fragment	HSCDC4A(E)	49/316
YCR36w	333	Ribokinase (other prokaryotic sugar kinases)	RBSK_ECOLI	38/96
YCR47c	275	ApoMet-methyltransferases	GLMT_RAT	26/301
YCR64c	136	Carboxypeptidases N	CBP8_HUMAN	27/88
		dipeptidyl-peptidase IV	DPP_LACLA	25/105
YCR69w (YCR70w)*	170	Peptidyl-prolyl-*cis-trans* isomerases	CYPH_CANAL	37/122
YCR72c	541	G-protein beta subunits	PR04_YEAST	23/278
			TUP1_YEAST	32/110
YCR98c	518	Sugar transporter/symporter	A40260(P)	25/179
YCR104w	124	Glucose repressor/cold shock inducible	SRP1_YEAST	27/115
			SCTIPI(E)	27/99
YCR106w	832	Gal4-like DNA/Zn binding domain	CYP1_YEAST	45/47
			GAL4_YEAST	19/168

ORF, predicted open reading frame (Oliver *et al.*, 1992); Family, functional protein family; Closest, sequence database identifier of the closest relative(s) from SWISS-PROT (default), PIR (P) or EMBL (E); %id/len, per cent amino-acid identity/length of the alignment.

*When the two adjacent ORFs YCR69w and YCR70w are cut and fused, correcting a possible frameshift sequencing error, they together represent a single member of the proline *cis-trans* isomerase family.

Pyroelectric X-ray generator

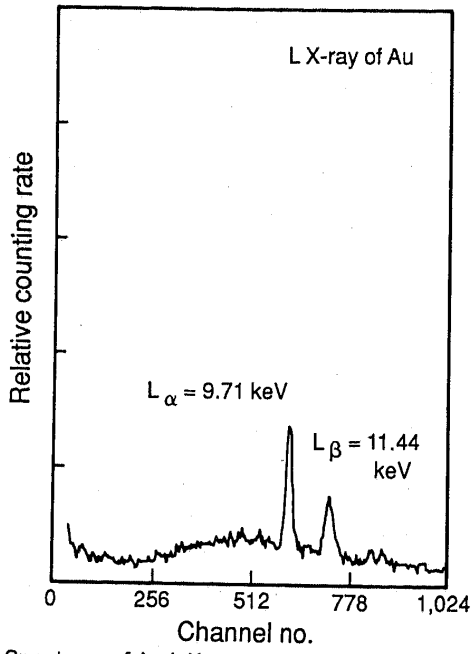
SIR — Pyroelectricity is a phenomenon analogous to the more familiar piezoelectricity, a transient voltage produced in response to a change in temperature rather than to a physical stress. Some pyroelectric crystals will also emit electrons[1–7]; electron energies up to 10^5 eV and electric fields as high as 10^6 V cm^{-1} have been obtained in single $LiNbO_3$ crystals. By placing a gold foil close to a pyroelectric crystal of $CsNO_3$, we have made a small X-ray generator which creates photons of about 20 keV.

When $CsNO_3$ is cooled from approximately 300 to 77 K, it produces an electric dipole, which can be maintained at constant temperature because the relaxation time is long[8]. During the cooldown, electrons are ejected from the negative end of the dipole. As the temperature is raised from 77 to about 150 K the dipole begins ejecting electron clusters in bursts from the positive end of the dipole. On warming, electron ejection starts and ends within a specific narrow temperature range for each crystal when the ambient vacuum and rate of change of the temperature is constant.

The sign of the ejected particles was determined to be negative by their deflection in magnetic and electric fields. The particles were detected with a silicon charged particle radiation detector and appeared to have high energy, in the MeV range. However, the beam of particles was attenuated to zero with 1.6 mg cm^{-2} mylar between the crystal and the detector.

When a cluster of electrons with a spread in time less than the resolving time of the system arrives at the detector, the output will reflect the total energy deposited in the detector by the electrons. Therefore, a closely packed cluster of about 15-keV electrons and a 5-MeV α-particle may appear to deposit the same amount of energy in the detector. Using α-emitting sources to calibrate the system, clusters of electrons with an apparent total energy as high as 10 MeV are observed.

In the figure we show the characteristic L X-rays of gold that are produced when a gold foil is adjacent to the positive end of a $CsNO_3$ crystal. A Si(Li) X-ray spectrometer with a 160-eV resolution was used to make this spectrum. These X-rays are produced as the temperature of the crystal is raised from 77 to 300 K and the foil is bombarded with ejected electrons. Fluorescent L X-rays result from vacancies produced in the L shell of Au atoms, which requires an energy of 14.3 keV. However, the K X-rays in Ag, requiring 25.5 keV, were



Spectrum of Au L X-ray detected with a Si(Li) X-ray spectrometer.

not produced, thus providing an upper limit for the energy of the ejected electrons.

The X-ray generator is a small vacuum cryostat with a liquid nitrogen reservoir, a thin window and a heater/cold finger. One end of the heater/cold finger is connected to a liquid nitrogen reservoir and the other faces the window. An ambient vacuum of 5×10^{-5} torr is sufficient. The crystal is placed on the end of the heater/cold finger with the positive end approximately perpendicular to the target and parallel to the window of the cryostat. A simple spring clip holds the crystal in place. The heater is a 47-Ω wire-wound resistor inside a brass tube that is the cold finger. The temperature of the cold finger is monitored with a thermocouple and the cryostat window is 4.5 mg cm^{-2} aluminium.

A typical $CsNO_3$ crystal grown in an aqueous solution is $0.2 \times 0.3 \times 0.8$ cm and is very fragile. When these crystals are cooled for the first time, they fracture and fragment. To prevent separation after fracturing, crystals are covered with a thin layer of epoxy. The ends may or may not be covered, as a thin layer of epoxy on the ends has little effect on the production of X-rays. The electrons accelerated by the crystal come from the ionization of molecules of the residual gases in the system.

The polarity of each crystal is determined by suspending it in liquid nitrogen in an electric field of known polarity and observing the alignment of the crystal with the electric field as it cools. If any part of the crystal is not covered with epoxy, all water must be removed immediately following this procedure.

To produce X-rays the heater is turned on and off, cycling the crystal between approximately 77 and 273 K. The cycle time is about 5 min, with about 1 min of X-ray production. With a thin window NaI(Tl) detector and a gold foil 0.5 cm from the crystal, 150,000 photons are detected in less than 1 min. No apparent reduction of X-rays is observed after repeated temperature cycles or if the crystal is maintained for long periods at either 77 or 300 K.

James D. Brownridge
Department of Physics, Applied Physics, and Astronomy,
State University of New York,
PO Box 6000, Binghamton,
New York 13902, USA

1. Rosenblum, B., Braunlich, P. & Carrico, J. P. *Appl. Phys. Lett.* **25**, 1 (1974).
2. Kalent'ev, V. A., Kortov, S. V. & Zatespin, A. F. *Sov. Tech. Phys. Lett.* **13**(11) (1987).
3. Rosenman, G. I., Rez, I. S., Chepelev, Yu. L. & Angert, N. B. *Sov. Phys. Tech. Phys.* **26**(2) (1981).
4. Kortov, V. S. *et al. Sov. Phys. Tech. Phys.* **25**(9) (1980).
5. Minakova, E. V., Tirkhomirova, N. A. & Khrustalev, Yu. A. *Phys. Chem. Mech. Surf.* **5**(7) 1861–1864 (1990).
6. Rosenman, G. I., Pechorskii, V. I., Chepelev, Yu. L., Boikova, E. I. & Issakova, L. E. *Phys. Stat. Sol.* **120**, 667 (1983).
7. Biedrzycki, K. *Ferroelectrics* **119**, 33 (1991).
8. Brownridge, J. D., Telesca, A. J., Stannard, C. R. & O'Brien, T. P. *Phys. Tech.* **28**(7), 482 (1990).

APC mutations

SIR — The function of multimeric proteins, or multiprotein aggregates, can be altered by inhibitory polypeptides produced through the action of dominant negative mutations[1]. The discovery of loci involved in colon cancer[2] which are dominant for polyp formation prompted Bourne to propose that these mutations exert their phenotype by this mechanism[3]. However, in a more recent paper, Groden *et al.*[4] argued that the nature of the observed mutational events in these patients with adenomatous polyposis coli (APC), which were found to be nonsense mutations and frameshifts, is inconsistent with such a model (the observed mutational events are incompatible with a dominant negative phenotype).

The consequence of such mutations, however, need not only be the inactivation of the gene and its protein product; through the process of translational reinitiation they can also alter the function of the product. In bacteria[5], nonsense and frameshift events that create or activate barriers to translation result in the production of dominant restart polypeptides in genes or proteins that exhibit an appropriate structure and function. This process provides a ready explanation for the APC results, namely, the carboxy-terminal fragments so generated by reinitiated translation interfere with the normal functioning of APC and/or MCC (mutated in colon cancer) protein complexes. Indeed, such a phenomenon may be particularly important in the alteration of the function of proteins insensitive to amino-acid substitution. Such functions would be sensitive to various mutational events that alter the reading frame (single-base frameshifts, deletions, duplications and rearrangements) while providing only a limited target for missense mutations. These mutational events have been neglected of late because of the predominance of single-base substitutions in the oncogene/suppressor gene systems so far characterized.

Alasdair Gordon
Section de Biologie,
Institut Curie,
26 rue d'Ulm,
75231 Paris Cédex, France

1. Herskowitz, I. *Nature* **329**, 219–222 (1987).
2. Kinzler, K. W. *et al. Science* **251**, 1366–1370 (1991).
3. Bourne, H. R. *Nature* **351**, 188–190 (1991).
4. Groden, J. *et al. Cell* **66**, 589–600 (1991).
5. Sarabhai, A. & Brenner, S. *J. molec. Biol.* **27**, 145–162 (1967).

Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

Time machine from dust

SIR — Allen and Simon[1] explain how a pair of cosmic strings, if they exist, could form a time machine. However, there are several other space-time metrics of general relativity which have this property, some of them quite simple.

Probably the simplest is the one discovered long ago by van Stockum[2,3]. This refers to the gravitational field outside an infinitely long circular cylinder in steady, rigid rotation. The cylinder is made of dust, that is, matter whose constituent parts have no interaction except their mutual gravitation. The properties of the system are governed by the mass per unit length, m. When m is not very large the gravitational field outside the matter is just that of a static infinite cylinder. This is surprising because it means that, contrary to expectations from Mach's principle, the rotating matter does not drag the exterior space-time with it.

When m is large (in fact, greater than 2×10^{27} g cm^{-1}), the rotating matter does drag the exterior space-time, and closed time-like lines exist outside the cylinder. Therefore an observer, using rockets, could travel into his past: the