Combined Functional Genomic Maps of the C. elegans DNA Damage Response

Simon J. Boulton, *et al.*

Science **295**, 127 (2002);
DOI: 10.1126/science.1065986

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

6. M. Misago *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 11766 (1995).
7. M. Oh-eda *et al.*, *Eur. J. Biochem.* **268**, 1280 (2001).
8. D. R. P. Tulsiani, H. Yoshida-Komiya, Y. Araki, *Biol. Reprod.* **57**, 487 (1997).
9. K. E. Panter *et al.*, *J. Nat. Toxins* **8**, 53 (1999).
10. D. R. P. Tulsiani, T. M. Harris, O. Touster, *J. Biol. Chem.* **257**, 7936 (1982).
11. B. Henrissat, *Biochem. Soc. Trans.* **26**, 153 (1998).
12. K. W. Moremen, in *A Comprehensive Handbook*, B. Ernst, G. Hart, P. Sinay, Eds., (Wiley, New York, 2000), p. 81.
13. D. R. P. Tulsiani, M. D. Skudlarek, M.-C. Orgebin-Crist, *Biol. Reprod.* **42**, 843 (1990).
14. S. Hiramoto *et al. Biochem. Biophys. Res. Comm.* **241**, 439 (1997).
15. M. D. Skudlark, M.-C. Orgebin-Crist, D. R. P. Tulsiani, *Biochem. J.* **277**, 213 (1991).
16. D. R. P. Tulsiani, M. D. Skudlarek, M.-C. Orgebin-Crist, *J. Cell. Biol.* **109**, 1257 (1989).
17. S. A. Igdoura *et al. Eur. J. Cell Biol.* **78**, 441 (1999).
18. D. Chui *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1142 (2001).
19. N. Takahashi, N. Tomiya, in *Handbook of Endoglycosidases and Glycoamidases*, N. Takahashi, T. Muramatsu, Eds. (CRC Press, Boca Raton, FL, 1992), pp. 199–332.
20. H. Nakagawa *et al.*, *Anal. Biochem.* **226**, 138 (1995).
21. R. M. DePhilip, D. G. Danahey, *Biol. Reprod.* **37**, 1271 (1987).
22. S. Kohno, E. Ziparo, L. F. Marek, K. S. Tung, *J. Reprod. Immunol.* **5**, 339 (1983).
23. Although the reported adhesion assays were performed at 32°C, we performed the assay at 37°C, because small numbers of $MX^{-/-}$ spermatogenic cells adhered to Sertoli cells at 32°C presumably by an additional mechanism independent from MX defect.
24. T. O. Akama, M. N. Fukuda, unpublished observations.
25. D. J. Moloney *et al.*, *Nature* **406**, 369 (2000).
26. M. Demetriou, M. Granovsky, S. Quaggin, S. J. W. Dennis, *Nature* **409**, 733 (2001).
27. S. Tsuboi, M. Fukuda, *EMBO J.* **16**, 6364 (1997).
28. P. Maly *et al.*, *Cell* **86**, 643 (1996).
29. I. Stamenkovic, D. Sgroi, A. Aruffo, *Cell* **68**, 1003A (1992).
30. L. G. Baum *et al.*, *J. Exp. Med.* **181**, 877 (1995).
31. Available at www.gak.co.jp/FCCA.
32. The authors thank E. Ruoslahti, E. Adamson, Y. Yamaguchi, M. Fukuda, and E. Lamar; for critical readings of the manuscript; M. Dym and N. B. Hecht for suggestions; R. DePhilip for instruction for germ cell–Sertoli cell adhesion assay; M. Oh-eda, M. Misago, R. Aoki, S. Saburi, and J. Zhang for discussions; and A. Pai for technical assistance. The present study was supported by NIH grants CA71932 (to M.N.F.), GM47533 and RR05351 (to K.W.M.), CA 42595 and HD05863 (to J.L.M.), and BCRP 4KB-0106 (to S.N.).

21 August 2001; accepted 1 November 2001

Combined Functional Genomic Maps of the *C. elegans* DNA Damage Response

Simon J. Boulton,[1,2] Anton Gartner,[3] Jérôme Reboul,[1] Philippe Vaglio,[1] Nick Dyson,[2] David E. Hill,[1] Marc Vidal[1]*

Many human cancers originate from defects in the DNA damage response (DDR). Although much is known about this process, it is likely that additional DDR genes remain to be discovered. To identify such genes, we used a strategy that combines protein-protein interaction mapping and large-scale phenotypic analysis in *Caenorhabditis elegans*. Together, these approaches identified 12 worm DDR orthologs and 11 novel DDR genes. One of these is the putative ortholog of hBCL3, a gene frequently altered in chronic lymphocytic leukemia. Thus, the combination of functional genomic mapping approaches in model organisms may facilitate the identification and characterization of genes involved in cancer and, perhaps, other human diseases.

Inherited cancer predisposition syndromes such as Li-Fraumeni syndrome, xeroderma pigmentosum, and hereditary nonpolyposis colon cancer result from defects in DNA repair or DNA damage checkpoint pathways (collectively referred to as the DDR) (*1*). In wild-type cells, checkpoint pathways induce a transient cell-cycle arrest in response to DNA damage, thus providing the necessary time for DNA repair to occur, and a variety of DNA repair pathways correct the various types of DNA lesions (*2*). Alternatively, in metazoan organisms, checkpoint pathways can also induce apoptosis, thereby eliminating compromised cells (*3*). *Caenorhabditis elegans* is the simplest metazoan model organism that can be used to study the DDR (*4*, *5*). After DNA damage, checkpoint pathways induce cell-cycle arrest or apoptosis of mitotic or pachytene cells of the adult germ line, respectively. These two cell types are located in spatially distinct regions.

To identify novel *C. elegans* DDR genes, we used a combination of functional genomic mapping approaches. High-throughput (HT) methods such as transcription profiling, protein interaction mapping, and large-scale phenotypic analysis have been applied individually to worm biology with considerable success (*6*, *7*). Consequently, hypotheses of function are now available for hundreds of previously uncharacterized genes. Although no single HT method can unequivocally define gene function, combining the data obtained from any of these complementary approaches is likely to provide greater functional insight (*7*). Here we have chosen to combine protein-protein interaction mapping and HT phenotypic analysis for the following three reasons. Because the function of most known DDR proteins is based on their ability to mediate protein-protein interactions, putative *C. elegans* DDR orthologs were used to generate a DDR protein interaction map. To demonstrate biological relevance, defects in the DDR were then analyzed systematically for each of the corresponding genes after HT RNA-mediated interference (RNAi). Lastly, in addition to potentially identifying novel DDR genes, combining these two approaches has the advantage of providing potential insights into how the corresponding proteins function, on the basis of the identity of their interacting partners.

To date, only three genes (*mrt-2*, *rad-51*, and *mre-11*) have been experimentally implicated in the *C. elegans* DDR (*4*, *5*, *8*). To identify potential worm DDR orthologs, we used known DDR proteins, including those implicated in nucleotide excision repair (NER), mismatch repair (MR), base excision repair (BER), nonhomologous end joining (NHEJ), homologous recombination (HR), and checkpoint pathways, as to query the *C. elegans* predicted proteome by BLAST [see legend to Web table 1 (*9*)]. A total of 75 putative *C. elegans* DDR orthologous open reading frames (dORFs) were identified and cloned using the Gateway recombinational cloning system (*10*) (Web table 1). The dORFs were then transferred to two-hybrid destination vectors to express either DNA binding domain (DB) fusions (DB-dORFs) or activation domain (AD) fusions (AD-dORFs) for protein-protein interaction analysis (*11*).

Conserved interactions, or interologs (*12*), were anticipated among worm DDR proteins on the basis of interactions reported between their potential orthologs in other organisms. To identify such DDR interologs, we first tested all possible pairwise combinations between the 75 DB-dORFs and AD-dORFs in a matrix setting. Among 33 putative interologs, 17 scored positive in the two-hybrid matrix (Table 1 and Web fig. 1). This represents a 51% success rate in detecting predicted protein interactions using this yeast two-hybrid system, which is in the same range as has been described previously for other protein interaction mapping projects (*13*). If one assumes that conservation of interaction between two putative orthologs is a reasonable indication of functional conservation, our search for interologs would suggest that at least 30 predicted worm DDR proteins are bona fide orthologs.

[1]Dana-Farber Cancer Institute and Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [2]Massachusetts General Hospital Cancer Center, Charlestown, MA 02129, USA. [3]Max Planck Institute for Biochemistry, D 82152 Martinsried, Germany.

*To whom correspondence should be addressed. E-mail: marc_vidal@dfci.harvard.edu

The matrix experiment also detected novel potential interactions (Table 1 and Web fig. 1). For example, MRE-11 and the *C. elegans* ortholog of poly-ADP-ribose polymerase (PRP-1) were found to interact in the yeast two-hybrid matrix. Both proteins have been implicated in telomeric maintenance in other organisms (*14, 15*). Thus, the MRE-11/PRP-1 interaction might occur in vivo at telomeres. The interactions between KIN-20, ATL-1 (ATR-like), BLM-1 (ortholog of Bloom's syndrome protein), and PCN-1 also suggest previously unrecognized connections between DDR proteins (Table 1 and Web fig. 1). ScHrr25, the budding yeast ortholog of KIN-20, is a protein kinase required for the DDR (*16*). *hrr25* mutants exhibit a radiation-sensitive phenotype and show defects in the transcriptional response to DNA damage after hydroxyurea treatment, whereas ScHrr25 has been shown to localize to sites of double-strand breaks (DSBs) (*16*). Perhaps these proteins, and potentially their respective human orthologs, are involved in a common DDR signaling pathway.

To identify new potential *C. elegans* genes involved in the DDR, proteome-wide two-hybrid screens were performed with 67 DB-dORF protein fusions [eight DB-dORFs scored as self-activators and, therefore, were not amenable to two-hybrid screening (*17*)]. No detectable interactions were found for 22 (32%) DB-dORFs. For the remaining 45 DB-dORF baits, a total of 165 interacting sequences, or interaction sequence tags (ISTs), were recovered, of which 125 are predicted to encode novel proteins (Web table 2). In addition, the proteome-wide screens recovered many of the interologs identified in the matrix experiment and also identified four more interologs (Table 1 and Web table 2).

To visualize the IST data, a protein-protein interaction map was constructed with the spring layout algorithm from "Algorithms for Graphic Drawing" (Fig. 1 and Web fig. 2). As observed previously, many two-hybrid interactions form IST clusters (*11*), which are defined as contiguous two-hybrid connections that form closed loops (for example, X/Y/Z/. . .n. . .X). They have been proposed to increase the likelihood of biological relevance for the corresponding potential interactions (*11*). A number of putative DNA damage checkpoint orthologs are linked by overlapping IST clusters in the DDR map: for example, ATL-1/PCN-1/MRT-2/HUS-1/F56D12.5/COH-2/ATL-1, ATL-1/BLM-1/KIN-20/PCN-1/ATL-1, and HUS-1/PDI-2/MRT-2HUS-1 (Fig. 1 and Web fig. 2). The worm orthologs of checkpoint proteins group together in one discrete region of the DDR map, whereas the repair proteins are grouped in other distinct regions (Fig. 1 and Web fig. 2). This reinforces the idea that two-hybrid protein interaction mapping tends to recapitulate actual connections. In addition, a number of novel proteins seem to link checkpoint IST clusters and groups of DNA repair proteins. This suggests that seemingly distinct DDR pathways might be physically linked. For example, COH-2, related to the SpRad21 cohesin protein that is implicated in DSB repair, is potentially associated with checkpoint proteins in two ways. First, COH-2 can interact directly with ATL-1. Second, it can also interact with the predicted protein F56D12.5, which itself can interact with the checkpoint protein HUS-1(ScMec3/SpHus1) (Fig. 1).

Although the yeast two-hybrid system used here minimizes the rate of false positives (*13*), the interactions identified should be considered merely as biological hypotheses until validated in vivo. Therefore, to determine which of the DDR orthologs and their potential interactors are in fact required for DDR processes in vivo, we subjected each gene to "RNAi by feeding" (*18*). We chose to focus our efforts specifically on the response to γ-irradiation by scoring for the following four DDR phenotypes: (i) defect in mitotic cell cycle arrest (Cca), (ii) reduction of apoptosis of pachytene nuclei (Rap), (iii) increase in apoptosis of pachytene nuclei (Iap), and (iv) radiation sensitivity of progeny (Rad) (Web fig. 3).

RNAi by feeding experiments conferred DDR phenotypes after γ-irradiation in each of three separate experiments for a total of 23 ORFs [RNAi did not confer any detectable or reproducible DDR phenotype for 169 ORFs after γ-irradiation (*19*)]. The newly identified DDR ORFs could be organized into two distinct phenotypic classes or "phenoclusters" based on similarity of their RNAi phenotype (Figs. 1 and 2 and Web figs. 4 and 5). One phenocluster included 11 ORFs that conferred similar defects to *mrt-2(RNAi)* animals, i.e., a Cca, a Rap, and a Rad phenotype. The other phenocluster included 12 ORFs that conferred a phenotype similar to that observed following RNAi of *mre-11*, i.e., an Iap and a Rad phenotype. The two phenoclusters contain both worm orthologs of known DDR genes and a number of their novel potential interactors (Figs. 1 and 2).

Defects in genes involved in checkpoint pathways would be expected to confer a Cca, a Rap, and a Rad phenotype. Indeed, RNAi of the putative checkpoint genes *hus-1*, *hpr-9*, *hpr-17*, *hsr-9*, *atm-1*, and *atl-1* conferred this phenotype (Fig. 2). These defects were all qualitatively similar to that observed for *mrt-2* (Web fig. 3), although in each case the penetrance of the phenotype varied somewhat. These phenotypes, taken together with interologs such as MRT-2/HPR-9, MRT-2/HUS-1, and HPR-17/RFC-4, strongly suggest that these DNA damage checkpoint genes are functionally conserved in *C. elegans*.

Checkpoint defects were also observed for five ORFs not previously implicated in this function. Those correspond to *pdi-2*, *exo-3*, *rfc-4*, *C04F12.3*, and *ubc-9* (Figs. 1 and 2 and Web figs. 4 and 5). *pdi-2* and *rfc-4* appear to perform essential functions because RNAi of these genes resulted in sterile (Ste) and dumpy (Dpy) animals and in embryonic lethality (Emb), respectively. However, the phenotypes in each case were only partially penetrant and "escapers" displayed defects in the DNA damage checkpoint (Figs. 1 and 2 and Web figs. 4 and 5). *pdi-2* encodes a β-subunit of protein disulfide isomerase that was found to associate with MRT-2 and HUS-1 checkpoint proteins. How PDI-2 participates in the checkpoint process is unknown, although its ability to associate with

Table 1. Interactions detected between worm orthologs of known DDR proteins. Putative worm DDR interologs were detected in the matrix (upper left corner), detected in proteome-wide library screens (lower right corner), or not detected at all (upper right corner). In addition, novel potential interactions that were also detected between DDR orthologs in the proteome-wide library screens (lower left corner).

Interologs DETECTED

H. sapiens			*C. elegans*		
MSH2	+	MSH6	H26D21.2	+	Y47G6A.11
XPA	+	RAD23	K07G5.2	+	ZK20.3
XPA	+	XPF1	K07G5.2	+	C47D12.8
XPF	+	ERCC1	C47D12.8	+	F10G8.7
UBE2	+	UBE2N	C35B1.1	+	F39B2.2
UBE2V2	+	UBE2N	Y94H6A_148.J	+	F39B2.2
RAD54	+	RAD51	W06D4.6	+	Y43C5A.6
RAD51	+	RAD51	Y43C5A.6	+	Y43C5A.6
RAD50	+	MRE11	T04H1.4	+	ZC302.1
KU70	+	KU80	Y47D3A.G	+	R07E5.8
RAD1	+	HUS1	Y41C4A.14	+	H26D21.1
RAD1	+	RAD9	Y41C4A.14	+	HPR-9
RAD17	+	RFC4	F32A11.2	+	F31E3.3
PCNA	+	PCNA	W03D2.4	+	W03D2.4
RFC1	+	PCNA	C54G10.2	+	W03D2.4
RFC1	+	RFC4	C54G10.2	+	F31E3.3
RFC1	+	RFC3	C54G10.2	+	C39E9.13

Interologs NOT DETECTED

H. sapiens			*C. elegans*		
MLH1	+	PMS2	T28A8.7	+	H12C20.2A
MSH2	+	FEN1	H26D21.2	+	Y47G6A.8
XPA	+	MSH2	H26D21.2	+	K07G5.2
XPF	+	MSH2	H26D21.2	+	C47D12.8
MSH2	+	PCNA	H26D21.2	+	W03D2.4
MSH2	+	ERCC5	H26D21.2	+	F57B10.6
RAD23	+	RAD7	ZK20.3	+	C02F5.7
XPA	+	MSH2	K07G5.2	+	H26D21.2
FEN1	+	PCNA	W03D2.4	+	Y47G6A.8
WRN1	+	KU70	F18C5.2	+	Y47D3A.G
RAD1	+	RAD17	Y41C4A.14	+	F32A11.2
PRIM1	+	PRIM2	F58A4.4	+	W02D9.1
RFC1	+	RFC2	C54G10.2	+	F58F6.4
RFC2	+	RFC3	F58F6.4	+	C39E9.13
RFC2	+	RFC4	F58F6.4	+	F31E3.3
RFC3	+	RFC4	C39E9.13	+	F31E3.3

NOVEL interaction DETECTED

H. sapiens			*C. elegans*		
RAD1	+	PCNA	Y41C4A.14	+	W03D2.4
ATR	+	PCNA	T06E4.3	+	W03D2.4
ATR	+	BLM1	T06E4.3	+	T04A11.6
CSNK1D	+	PCNA	F46F2.2	+	W03D2.4
CSNK1D	+	BLM1	F46F2.2	+	T04A11.6
CSNK1D	+	ATR	F46F2.2	+	T06E4.3
MRE11	+	PARP	ZC302.1	+	AC8.1
RAD51	+	RAD51L3	Y43C5A.6	+	C30A5.2

Interologs DETECTED in the library screens

H. sapiens			*C. elegans*		
BLM	+	TOP2A	T04A11.6	+	K12D12.1
MLH1	+	PMS2	T28A8.7	+	H12C20.2A
RAD23	+	PMSD4	ZK20.3	+	B0205.3
RAD51	+	UBE2I	Y43C5A.6	+	F29B9.6

known checkpoint proteins provides a starting point for more detailed molecular analysis. RFC-4, a replication factor C subunit, was also found to interact in the yeast-two hybrid with the checkpoint protein, HPR-17 (Fig. 1). Our observations suggest a model in which RFC-4 functions both in DNA replication and in a DNA damage checkpoint pathway.

exo-3 encodes the *C. elegans* ortholog of human APEX/REF1, a bifunctional enzyme that has been implicated previously in both DNA excision repair and a redox-dependent activity that stimulates transcription factor binding (*20*). We have uncovered an additional role for this enzyme in the DNA damage checkpoint. It is possible that EXO-3 may perform its role in the checkpoint at the level of damage detection or may be required for the induction of gene transcription necessary to elicit cell cycle arrest or apoptosis in response to damage.

A novel ORF implicated in the DNA damage checkpoint, *C04F12.3*, encodes a putative ortholog of human BCL3 that was found as an interactor with the checkpoint protein MRT-2 (Figs. 1 and 2). The *hBCL3* gene is frequently associated with chromosomal translocations and amplification in cancer and has been implicated in cell cycle control (*21*). If human BCL3 is also required for checkpoint integrity, a defect in this function could explain the effects of hBCL3 translocation and amplification in chronic lymphocytic leukemias.

Defects in genes required for DNA repair were expected to result in an Iap and Rad phenotype after γ-irradiation. Indeed, RNAi of the putative DNA repair genes *coh-2*, *rad-50*, *rad-51*, and *rad-54* conferred the phenotype expected for DSB repair defects (Figs. 1 and 2 and Web figs. 4 and 5). These phenotypes, together with the detection of RAD-54/RAD-51 and RAD-50/MRE-11 interologs (Figs. 1 and 2), demonstrate that these genes are required for DSB repair after γ-irradiation. In addition, we identified six novel genes that displayed an Iap and Rad phenotype after RNAi, similar to the defects observed in *mre-11(RNAi)* animals (Fig. 2 and Web fig. 3).

Two of the novel DNA repair genes encode a RAD-54 interactor (C02F12.4) and a RAD-51 interactor (Y116A8C.13), respectively. C02F12.4 contains a RhoGEF domain and Y116A8C.13 is a RAD-54 related protein (Web fig. 4). In addition, three novel DNA repair genes encode COH-2 interactors. F11E6.1 and F56D12.5 have no assigned function and F13D12.6 resembles serine carboxypeptidases (Web fig. 4). In addition to being able to interact with COH-2, F56D12.5 also interacts with the checkpoint protein HUS-1 (Figs. 1 and 2). The F56D12.5(RNAi) phenotype further suggests that DSB repair and DNA damage checkpoint

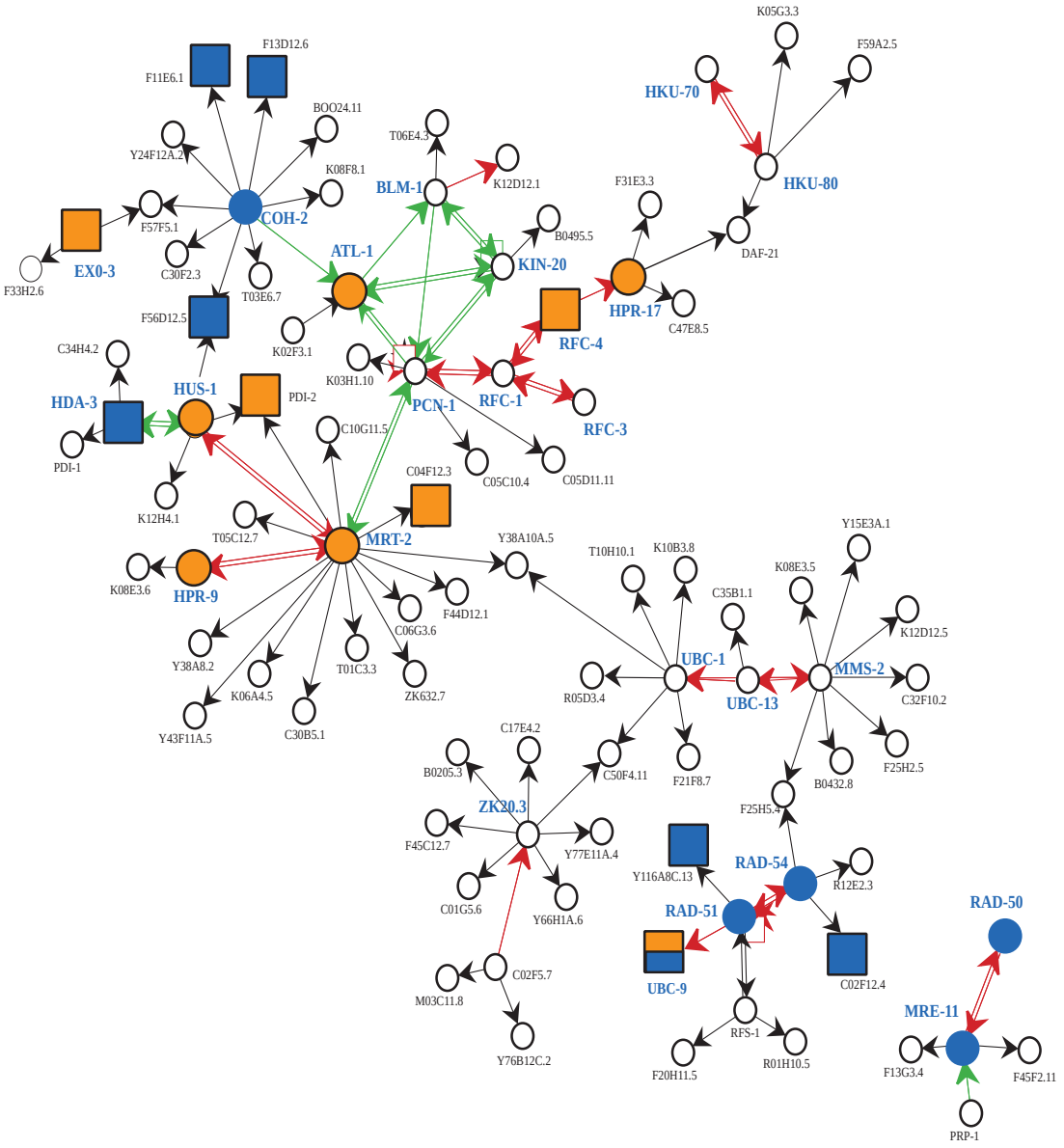

Fig. 1. DDR phenotypes overlap with the protein interaction map. Shown is part of the DDR protein interaction map derived from two-hybrid interactions detected in the matrix and in proteome-wide screens (see Web fig. 5 for the complete map). Arrows and circles (or squares) represent two-hybrid interactions and proteins, respectively. Arrows point directionally from the baits to the potential interactors. The color-code for interactions are as follows: red arrows, detected interologs; green arrows, novel interactions between DDR proteins; black arrows, interactions between DDR and novel proteins. Genes that give a similar DDR phenotype after RNAi encode proteins that tend to cluster on the protein interaction map. In this map, colored circles (orthologs of known DDR proteins) and squares (novel DDR proteins) indicate products of genes from the DNA repair and checkpoint phenoclusters shown in Fig. 2. The color code for the DDR phenotypes are as follows: orange, checkpoint defective; blue, DNA repair defective; blue and orange, checkpoint and repair defective.

pathways may be physically linked.

hda-3 is required for DNA repair and is predicted to encode a histone deacetylase (HDAC) that can interact with HUS-1 (Figs. 1 and 2). The human HUS-1 protein can also associate with a histone deacetylase, hHDAC1 (*22*). However, the biological relevance of this interaction has remained somewhat unclear. Our data suggest that *hda-3* is likely to be required for DSB repair and survival after γ-irradiation.

Lastly, we identified *ubc-9* as a gene that seems to be required for both DNA repair and DNA damage checkpoint function. *ubc-9(RNAi)* confers a Ste phenotype that was not fully penetrant. The progeny that survived displayed a checkpoint defect of mitotic nuclei as well as Iap and Rad phenotypes, consistent with a DNA repair defect. Accordingly, the yeast ortholog of UBC-9, a sumo-conjugating enzyme, is required for both DNA repair and normal mitotic cell division (*23*). Targets of Ubc9-dependent sumolation have been reported to function in many different cellular pathways in yeast and in mammalian cells (*24*). Our observations raise the possibility that components of both DNA repair and checkpoint pathways could be targets for UBC-9 dependent sumolation.

In summary, by combining large-scale protein interaction mapping and HT phenotypic analysis, we have identified 23 genes required for the DDR in *C. elegans*, 11 of which were not previously identified in any system. These data also highlight the extraordinary level of conservation of molecular mechanisms in DDR pathways. Of the new DDR genes that have apparent orthologs in human (10/11), it will be important to determine whether mutations in any of these genes give rise to cancer predisposition (Web fig. 4).

Among the potential interactors identified by protein interaction mapping, the hit rate of genes (11/125) that gave rise to a detectable DDR RNAi phenotype was strikingly higher than would be expected from random. This strongly suggests that combining protein interaction mapping and HT RNAi data can be used synergistically to identify novel genes involved in biological processes of interest. However, it is also important to consider why the remaining genes (90%) did not give rise to any detectable DDR phenotype (*19*). Although the phenotypes assayed under our conditions are likely to uncover genes involved in the sensing and repair of DSBs, they are unlikely to identify genes required for other DDR processes such as base excision, nucleotide excision, and mismatch repair pathways. Thus, we anticipate that future studies, using assays that measure sensitivity to other DNA damaging agents, such as UV or alkylating agents, are likely to reveal additional new DDR genes in the protein interaction map. Furthermore, alternative methods for detecting protein interactions and data sets from protein localization and transcription profiling experiments are also likely to further strengthen and expand the DDR map presented here (*7*). For example, transcriptional profiling is likely to identify new genes whose expression is induced in response to DNA damage.

Combining data from complementary large-scale approaches establishes a new paradigm in the field of functional genomics that makes it possible to greatly accelerate functional discovery. Importantly, this strategy should be applicable to other biological processes and/or model organisms. With a draft of the human genome in hand (*25*, *26*) and the recent description of RNAi effects in mammalian cells (*27*), this approach might soon be applicable to human biology.

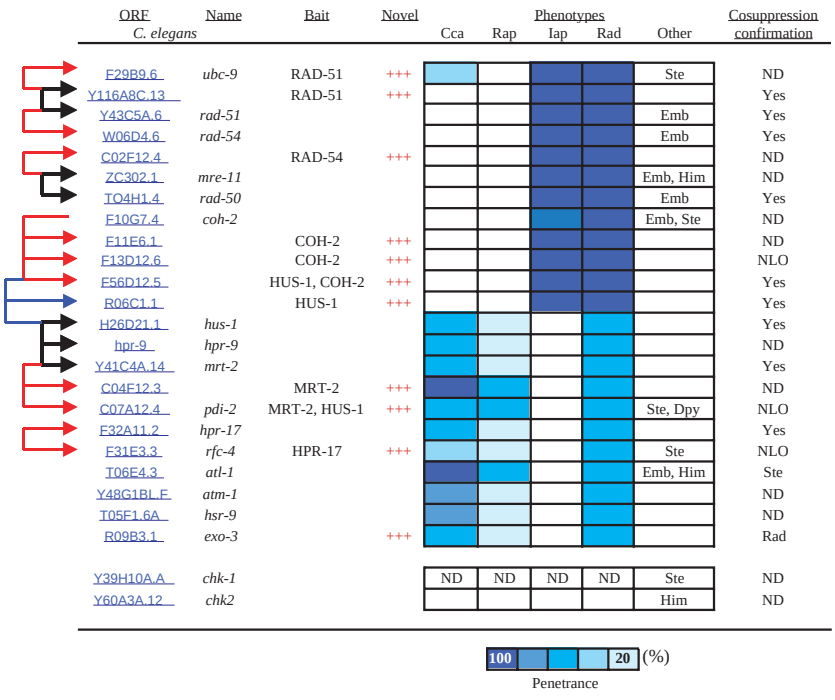

Fig. 2. Phenoclusters correlate with two-hybrid interactions. Four DDR phenotypes, Cca, Rap, Iap, and Rad, were scored after RNAi treatment for each one of 192 ORFs. ORFs that scored in similar ways for one or more of these four phenotypes were clustered together. Blue intensities relate to the level of penetrance for each phenotype (scale is indicated below with the percent of animals showing the defect). In some cases, the RNAi phenotypes were confirmed by co-suppression analysis (*31*) (NLO, no line obtained; ND, not determined). Additional phenotypes detected are indicated in the "Other" column (Ste, sterile; Emb, embryonic lethal; Him, high incidence of males; Dpy, dumpy). Genes belonging to either the checkpoint or repair "phenoclusters" encode proteins that tend to interact with each other. For each novel DDR gene (see red plus signs), the bait(s) used for their two-hybrid identification is shown in the "Bait" column. The arrows on the left show the two-hybrid interactions between these gene products with the color code as follows: black, interologs; red, novel interactions between the products of conserved and new DDR genes that belong to a common phenocluster; blue, novel interactions between the products of genes that belong to different phenoclusters.

References and Notes

1. L. Hartwell, T. Weinert, L. Kadyk, B. Garvik, *Cold Spring Harbor Symp. Quant. Biol.* **59**, 259 (1994).
2. T. Lindahl, R. D. Wood, *Science* **286**, 1897 (1999).
3. B. B. Zhou, S. J. Elledge, *Nature* **408**, 433 (2000).
4. S. Ahmed, J. Hodgkin, *Nature* **403**, 159 (2000).
5. A. Gartner, S. Milstein, S. Ahmed, J. Hodgkin, M. O. Hengartner, *Mol. Cell* **5**, 435 (2000).
6. P. W. Sternberg, *Cell* **105**, 173 (2001).
7. M. Vidal, *Cell* **104**, 333 (2001).
8. G. M. Chin, A. M. Villeneuve, *Genes Dev.* **15**, 522 (2001).
9. Supplemental data are available at *Science* Online at www.sciencemag.org/cgi/content/full/295/5552/127/DC1.
10. A. J. M. Walhout *et al.*, *Methods Enzymol.* **328**, 575 (2000).
11. A. J. M. Walhout *et al.*, *Science* **287**, 116 (2000).
12. L. Matthews *et al.*, *Genome Res.* **12**, 120 (2001).
13. The two-hybrid system used here reduces the occurrence of false positives (spurious interactions that bear no biological significance) by lowering the expression levels of the two hybrid proteins and by using multiple reporter genes that utilize different promoters (*28*). False negatives are reported interactions that can not be readily detected in the two-hybrid system. The percentage of false negatives in the DDR map is in agreement with data obtained from signal transduction and proteasome interaction maps, estimated to be approximately 55% (*10*, *29*) [reviewed in (*30*)].
14. S. J. Boulton, S. P. Jackson, *EMBO J.* **17**, 1819 (1998).
15. F. d'Adda di Fagagna *et al.*, *Nature Genet.* **23**, 76 (1999).
16. M. F. Hoekstra *et al.*, *Science* **253**, 1031 (1991).
17. A. J. M. Walhout, M. Vidal, *Methods* **24**, 297 (2001).
18. A. Fire *et al.*, *Nature* **391**, 806 (1998).
19. It is possible that RNAi might fail to confer an effect on genes actually required for DDR processes. For example, RNAi is not efficient at detecting the function of neuronal genes. However, 82% (14/17) of genes predicted to be required for the response to γ-irradiation in *C. elegans*, on the basis of work in other organisms, were identified in our screen (RNAi of *hku-70*, *hku-80*, or *dnl-4* did not confer DDR phenotypes). Furthermore, a number of yeast DDR genes are essential (e.g., Sc*MEC1*

and Sc*LCD1*). Indeed, 20.7% (35/169) of the genes that did not confer any DDR phenotypes gave Emb or Ste phenotypes (data not shown).
20. S. Xanthoudakis, T. Curran, *EMBO J.* **11**, 653 (1992).
21. L. Michaux *et al.*, *Genes Chromosomes Cancer* **15**, 38 (1996).
22. R. L. Cai, Y. Yan-Neale, M. A. Cueto, H. Xu, D. Cohen, *J. Biol. Chem.* **275**, 27909 (2000).
23. F. al-Khodairy, T. Enoch, I. M. Hagan, A. M. Carr, *J. Cell Sci.* **108**, 475 (1995).
24. S. Muller, C. Hoege, G. Pyrowolakis, S. Jentsch, *Nature Rev. Mol. Cell Biol.* **2**, 202 (2001).
25. J. C. Venter *et al.*, *Science* **291**, 1304 (2001).
26. E. S. Lander *et al.*, *Nature* **409**, 860 (2001).
27. S. M. Elbashir *et al.*, *Nature* **411**, 494 (2001).
28. M. Vidal, P. Legrain, *Nucleic Acids Res.* **27**, 919 (1999).
29. A. Davy *et al.*, *EMBO Rep.* **2**, 821 (2001).
30. A. J. M. Walhout, M. Vidal, *Nature Rev. Mol. Cell. Biol.* **2**, 55 (2001).
31. A. F. Dernburg, J. Zalevsky, M. P. Colaiacovo, A. M. Villeneuve, *Genes Dev.* **14**, 1578 (2000).
32. We thank J. Reid, A. J. M. Walhout, J. Lamb, M. Tewari, R. Abraham, and S. van den Heuvel for comments on the manuscript; D. Haber for valuable information; members of the Vidal lab for support; and the *Caenorhabditis* Genetics Center (CGC) for providing worm strains. S.J.B. was supported by Human Frontiers Science Project and Tosteson Postdoctoral Fellowships. A.G. was supported by the Max Planck Society (E. Nigg) and by Deutsche Forschungsgemeinschaft grant 703/1-1. Supported by grants 5R01HG01715-02 (National Human Genome Research Institute), P01CA80111-02 and 7 R33 CA81658-02 (National Cancer Institute), and 232 (Merck Genome Research Institute) awarded to M.V.

5 September 2001; accepted 30 October 2001

Construction and Analysis of a Human-Chimpanzee Comparative Clone Map

Asao Fujiyama,[1,2*†] Hidemi Watanabe,[1*†] Atsushi Toyoda,[1*] Todd D. Taylor,[1*] Takehiko Itoh,[3*] Shih-Feng Tsai,[4,5*] Hong-Seog Park,[6*] Marie-Laure Yaspo,[7*] Hans Lehrach,[7] Zhu Chen,[8*] Gang Fu,[8*] Naruya Saitou,[2*] Kazutoyo Osoegawa,[9] Pieter J. de Jong,[9] Yumiko Suto,[10] Masahira Hattori,[1*†] Yoshiyuki Sakaki[1,11*†]

The recently released human genome sequences provide us with reference data to conduct comparative genomic research on primates, which will be important to understand what genetic information makes us human. Here we present a first-generation human-chimpanzee comparative genome map and its initial analysis. The map was constructed through paired alignment of 77,461 chimpanzee bacterial artificial chromosome end sequences with publicly available human genome sequences. We detected candidate positions, including two clusters on human chromosome 21 that suggest large, nonrandom regions of difference between the two genomes.

With the draft human genome sequences (*1*, *2*), comparative genomics has become a powerful approach to extract genetic information from large stretches of nucleotide sequences through identification of conserved regions that are most likely functionally important. Genomic information is also the most valuable resource for understanding the genetic differences between species, a basis for deciphering how genome information is processed into phenotypes. Because chimpanzees are our closest relatives, the differences between us are less than with any other species, yet these differences are more likely to be important. It has been estimated that the sequence identity between human and chimpanzee is within the range of 98 to 99% (*3*–*12*). Thus, comparisons between humans and chimpanzees are the most efficient and effective approach to understand what makes us human.

In this report we present the construction and analysis of a first-generation human-chimpanzee comparative genomic map based on the alignments of 77,461 chimpanzee bacterial artificial chromosome (BAC) end sequences (BESs) to human genomic sequences obtained from the public databases. To prepare the BESs, we used two independently prepared BAC libraries, PTB1 and RPCI-43 (Table 1) (*13*). Briefly, we sequenced 64,116 BAC clones (roughly 3.3 times coverage of the currently available human contiguous genomic sequence) that produced 114,421 valid BESs (*13*). The BESs were then aligned with the RefSeq human genome contigs [National Center for Biotechnology Information (NCBI)] through NCBI-BLAST. The number of BESs having an alignment longer than 50 base pairs (bp) with ≧90% identity was 77,461 (*13*). Out of this number, 49,160 BESs from 24,580 clones formed paired ends where each pair was derived from the same clone. Only one end could be successfully aligned from the remaining 28,301 clones. The remaining 36,960 BESs that were not mapped to the human genome were categorized into three different classes: (i) those corresponding to repeat sequences (1168

Table 1. Summary of BES readings and mapping.

		Number of BAC clones*	Number of BESs
Sequenced		64,116	114,421†
Mapped		52,881	77,461‡
	Paired ends§	24,580	49,160
	Singletons‖	28,301	28,301
Unmapped			36,960
	Repeats		1,168
	Human¶		20,376
	Nonhuman		515
	No hit		14,901

*These clones were selected from the PTB1 or RPCI-43 chimpanzee male BAC libraries, and both ends of the inserts were sequenced (53,240 BESs from PTB1 and 24,221 BESs from RPCI-43 are mapped). †Number of total BESs. ‡Number of mapped BESs. §Number of mapped BESs derived from the same BAC clone. Long BES pairs (>300 kb) were not used for the mapping. ‖Number of the singleton (lone) BESs mapped on the human genome sequence. ¶Contains hits to sequences not included in the NCBI contigs or hits to mRNAs only.

[1]RIKEN Genomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. [2]National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8450, Japan. [3]Mitsubishi Research Institute, 3-6, Otemachi 2-chome Chiyoda-ku, Tokyo 100-8141, Japan. [4]National Yang-Ming University, 128, Yen-chiu, Yuan Road, Sec. 2, Taipei 115, Taiwan. [5]National Health Research Institutes, 128, Yen-chiu, Yuan Road, Sec. 2, Taipei 115, Taiwan. [6]Genome Research Center, KRIBB, 52 Oun-dong, Yusong-gu, Daejeon 305-333, Korea. [7]Max Planck Institute for Molecular Genetics, Ihnestrasse 73, D-14195 Berlin-Dahlem, Germany. [8]Chinese National Human Genome Center at Shanghai, 351, Guo Shou Jing Road, Shanghai 201203, China. [9]Children's Hospital Oakland Research Institute, BACPAC Resources, 747 52nd Street, Oakland, CA 94609–1809, USA. [10]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Hongou, Bunkyou-ku, Tokyo 113-0033, Japan. [11]Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.

*These authors form the International Consortium for the Sequencing of Chimpanzee Chromosome 22.
†To whom correspondence should be addressed. E-mail: afujiyam@gsc.riken.go.jp, watanabe@gsc.riken.go.jp, hattori@gsc.riken.go.jp, sakaki@gsc.riken.go.jp