AD_____

Award Number: DAMD17-03-1-0438

TITLE: Studies of Neurofibromatosis-1 Modifier Genes

PRINCIPAL INVESTIGATOR: Andre Bernards, Ph.D.

CONTRACTING ORGANIZATION: Massachusetts General Hospital
Boston, MA 02114

REPORT DATE: June 2005

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-06-2005	Annual	1 Jun 2004 - 31 May 2005

4. TITLE AND SUBTITLE

Studies of Neurofibromatosis-1 Modifier Genes

5a. CONTRACT NUMBER

5b. GRANT NUMBER
DAMD17-03-1-0438

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Andre Bernards, Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Massachusetts General Hospital
Boston, MA 02114

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT – SEE ATTACHED PAGE

15. SUBJECT TERMS
Neurofibromatosis-1, Neurofibroma burden modifier genes

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	33	

Abstract

This project aims to collect NF1 patient DNAs required to identify neurofibroma burden modifier genes, to perform an allele association study for three classes of potential modifiers, and to evaluate more global approaches. Over four years we aim to collect 1200 DNAs from adult NF1 patients that represent the top and bottom 20% of dermal neurofibroma burden in various age cohorts. We will use these DNAs in a case-control allele association study that will test whether neurofibroma numerical variability reflects (1) allelic differences in genes that maintain genome stability; (2) differences in the *NF1* gene itself or in closely linked genes; or (3) differences in genes involved in signaling between neurofibroma constituent cells. In the second year of this project we have continued to make progress towards our patient recruitment goal, in part by enlisting additional clinical collaborators. Among other major progress, we have implemented a relational database of approximately 25% of human genes, and w have used this database to identify and prioritize approximately 1000 candidate neurofibroma burden modifiers. We have also begun the genotyping of selected high priority SNPs and have published three review papers describing our bioinformatics efforts.

Table of Contents

Introduction

Neurofibromatosis type 1 (NF1) affects approximately 2-3 in 10,000 worldwide. A high degree of unpredictability and variability of symptoms is among the hallmarks of NF1. This variable expressivity increases patient anxiety and is a serious problem for those conducting clinical trials. Thus, significant resources have been devoted to better describing the natural history of NF1, without much to show for the effort. We have chosen an alternative approach to increase the predictability of NF1, based on studies that have implicated symptom-specific modifier genes as important determinants of the clinical variability in NF1 (Easton et al., 1993; Szudek et al., 2000; Szudek et al., 2002; Szudek et al., 2003). Our specific aims are to collect somatic DNAs from 1200 NF1 patients that represent the top and bottom 20% of dermal neurofibroma burden and to perform a case-control allele association study to evaluate potential hypomorphic alleles of three classes of candidate neurofibroma burden modifiers. The classes of potential modifier genes that we will evaluate are (1) genes implicated in maintaining genome stability, based on the idea that loss of the wild-type *NF1* allele may be rate-limiting in tumor development, (2) genes included in so-called NF1 microdeletions, based on the observation that patients who carry microdeletions often develop large numbers of early onset neurofibromas, and (3) genes involved in signal transduction between the cell types that make up neurofibromas, based on the idea that cross-talk between neurofibroma constituent cells may be essential for tumor development.

Body

Our Statement of Work described the following aims for the first two years of this project:

Year 1: (**1**) Assembly of 500-600 member case-control patient panel; (**2**) Identify and prioritize Single Nucleotide Polymorphisms (SNPs) in all classes of potential modifiers; (**3**) Design and validate SNP genotyping assays; (**4**) Begin genotyping of the most common, >0.1 Variant Allele Frequency (VAF) SNPs in the 600 member case-control NF1 patient exploration panel.

Year 2: (**1**) Assembly of 750-900 member case-control patient panel; (**2**) Identify and prioritize additional SNPs in all classes of potential modifiers; (**3**) Design and validate SNP genotyping assays; (**4**) Continue genotyping of >0.1 VAF SNPs in exploration panel. (**5**) Evaluation of new genotyping technology

Patient recruitment progress (aim 1, years 1 and 2)

Obviously, the recruitment of sufficient numbers of carefully evaluated eligible patients is critically important for the eventual success of this project. A sufficient number during the first years of this project means 600 patients equally distributed over low and high neurofibroma burden categories. These 600 patients constitute our 'exploration panel', in which we will genotype the most common (>0.1 VAF) potentially hypomorphic SNPs. The exploration panel will be supplemented in later years by a independent 600 member 'validation panel', which will be used to confirm any detected allele association. The full 1200 member patient panel can also be used to detect statistically significant allele associations for less frequent SNPs.

In our first annual report we indicated that we had not included obtaining regulatory approval as a separate goal. We had not done this, because the Army Regulatory Compliance Office had previously approved our essentially identical pilot study, funded as an Idea Award in 1999. Thus, we did not anticipate it would take over 8 months to obtain approval, nor did we foresee that major changes in recruitment procedures and study design would be mandated. One problem was that the Army Compliance Office did not agree to new HIPAA-mandated language in our MGH/Partners Healthcare consent form. We did not have the authority to modify this admittedly convoluted language, and a conference call between all concerned authorities did nothing to resolve this issue. Another problem was that a previously approved method of recruiting patients who contacted us after learning about our study was no longer deemed allowable. In the end the only practical solution to both problems was our agreement to stop recruiting patients ourselves, and to henceforth restrict our analysis to de-identified DNA samples provided by collaborators.

Collaborator	Location	# DNAs available	# prospective patients
Evans, Gareth	Manchester, UK	0	150
Ferner, Rosalie	London, UK	0	>100
Lázaro, Conxi	Barcelona, Spain	55	30-60
Legius, Eric	Leuven, Belgium	0	>75
Mautner, Victor-Felix	Hamburg, Germany	288	300
Messiaen, Ludwine	Ghent, Belgium	50	50-70
Locally recruited	Boston, MA	64	100
Total		457	805-875

Table 1. Collaborators and the numbers of available and to-be-recruited patients. Taken from our original grant application.

Table 1, taken from our grant application, indicates the number of available and to-be-recruited patients that six collaborators had agreed to contribute. Our own Institutional Review Board (IRB) and the Army Compliance Office ruled that DNA samples from 457 previously recruited patients were exempt and could be analyzed immediately. Most samples in this category had been promised to us by a single collaborator, Dr. Victor Mautner from Hamburg, Germany. Unfortunately, as indicated in our previous report, upon delivery 116 of the German samples turned out to be from patients who did not meet our neurofibroma burden eligibility criteria, and a further 51 samples contained no detectable DNA. Thus, only 121 German samples, 55 Spanish samples given to us by Dr. Lazaro, and 64 DNAs from patients recruited by us during our previous pilot project were available for immediate analysis. Delivery of the 50 samples promised by Dr. Messiaen continues to be delayed after her recent move from Ghent, Belgium to Dr. Korf's Department at the University of Birmingham, AL. Thus, rather than the promised 457 samples, only 240 patient samples were available for immediate analysis at the start of this study. For all but the most common SNPs, this number is not sufficient to reach statistical significance, prompting us to devote much initial effort to identifying further clinical collaborators.

Another setback was that Dr. Mautner unexpectedly communicated his unwillingness to recruit further patients without receiving significant monetary compensation. Rather than requesting support for a full time position when our budget could still be modified, his demand for financial compensation came only after our proposal was funded. This experience is among those that have stimulated our efforts to enlist additional clinical collaborators. After presenting the design of this study as a platform presentation at the 2003 annual National Neurofibromatosis Foundation consortium meeting, and as posters at the 2004 and 2005 meetings, several potential new collaborators expressed interest in helping us achieve our goals. Enthusiasm was particularly stimulated by a keynote address by Dr. Arnold Levine at the 2005 meeting. In his presentation Dr. Levine spoke about his recent human genetic studies, which identified *MDM2* and *AKT1* SNPs associated with accelerated tumor development (Bond et al., 2004; Arva et al., 2005). After his talk several clinicians, including Drs. Bruce Korf, Susan Huson, Meena Upadhyaya and even Victor Mautner expressed renewed enthusiasm in being part of our NF1 modifier study.

As a result of our efforts to advertise this study we have thus far signed up two major new collaborators. Thus, Cynthia MacKenzie on behalf of Dr. James Tonsgard recently obtained local IRB approval to provide us with de-identified blood samples from eligible patients seen at the University of Chicago NF clinic. In last year's progress report we also mentioned that Dr. Susan Huson had expressed interest in serving as our study coordinator in Great Britain. Dr. Huson is the author of several clinical studies of NF1 (Huson and Hughes, 1994), and has had a long interest in the role of modifier genes in determining disease outcome. At the 2005 meeting of the recently renamed Children's Tumor Foundation (CTF), Dr. Huson continued to express her interest in coordinating patient recruitment at U.K. clinics. However, her efforts to obtain local regulatory approval have been slowed by her recent move from Oxford to Manchester. Dr. Huson's latest estimate is that she will obtain regulatory approval later this year, after which the process of getting her approved by the Army Regulatory Compliance Office can start. Separately, at the 2005 CTF meeting Dr. Meena Upadhyaya (Bristol, U.K) indicated her willingness to provide us with samples from already recruited eligible patients. Many of her patients were recruited before informed consent was mandatory, so we expect that these samples will be exempt. As we have done in the past, we will work with collaborators to facilitate their IRB approval and resolve other regulatory issues. However, in general it has proven a frustratingly low process to get collaborators approved.

Bioinformatics Progress:

The second aim for years 1 and 2 of this project was to identify and prioritize SNPs in all three classes of potential neurofibroma burden modifiers. In our original proposal we described the design of relational databases to collect and manage information on potential hypomorphic alleles among a comprehensive set of genome stability genes. In year 1 we essentially completed this goal for genome stability genes (category 1), identifying 964 missense SNPs among 319 genes. For statistical and other reasons explained in our original proposal, only missense SNPs with a VAF >4% will be analyzed. Among the 964 missense SNPs in genome stability genes, 176 have a VAF >4%. This compares with 576 missense SNPs (110 >4%) in 244 candidate modifier genes that we had identified when this proposal was written. In the first year of this project we also identified a comprehensive set of highly polymorphic variable nucleotide tandem repeats and SNPs in the 1.5-2.0 MB segment deleted in patients with microdeletions (category 2) (Jenne et al., 2001). Finally, in year 1 we began the process of identifying candidate genes involved in signal transduction between the cell types that make up neurofibromas (category 3). In last year's progress report we described that as part of this effort we had generated a second set of databases to collect information on human members of the Ras GTPase superfamily, their regulators and interactors/effectors. Using a cross-species BLAST approach, we identified 159 human Ras superfamily members, 174 potential and confirmed human Ras superfamily GTPase Activating Proteins (GAPs), 155 potential and confirmed guanine nucleotide exchange factors (GEFs), and 359 GTPase effectors/interacting proteins. The PI of this proposal presented a talk about this work at the 2004 FASEB meeting on small GTPases in Snowmass, CO, which led to an invitation by Drs. William Balcher, Channing Der and Alan Hall to write the introductory chapter for the latest edition of a three book set in the Methods in Enzymology series focusing on Ras-related GTPases (Bernards, 2005). This chapter, included

as Appendix 1, describes that roughly 3% of human genes code for Ras superfamily members, their regulators and effectors. The GTPase and Affiliated Protein database will be made available online at the time of publication. Based on our bioinformatics work we also wrote two reviews on the regulation of GTPase Activating Proteins (Bernards and Settleman, 2004), and on their roles in growth factor signaling (Bernards and Settleman, 2005). Both were published during the current funding period and are included as Appendices 2 and 3.

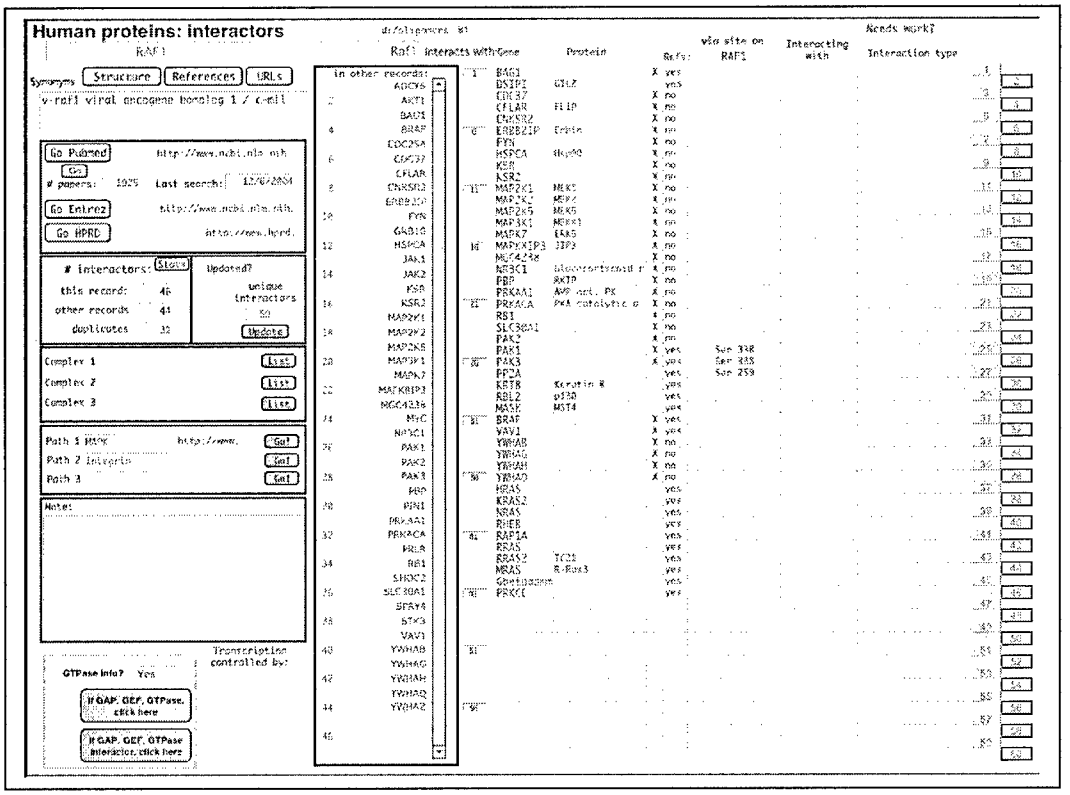


Figure 1. Human database layout showing information on protein interactors, in this example for the Raf1 kinase. Each gene-specific record can be viewed in 73 layouts, 60 of which are used to document protein-protein interactions. The remaining 13 summarize structural, functional, SNP-related, and reference information. See Appendix 1 for more information on database design.

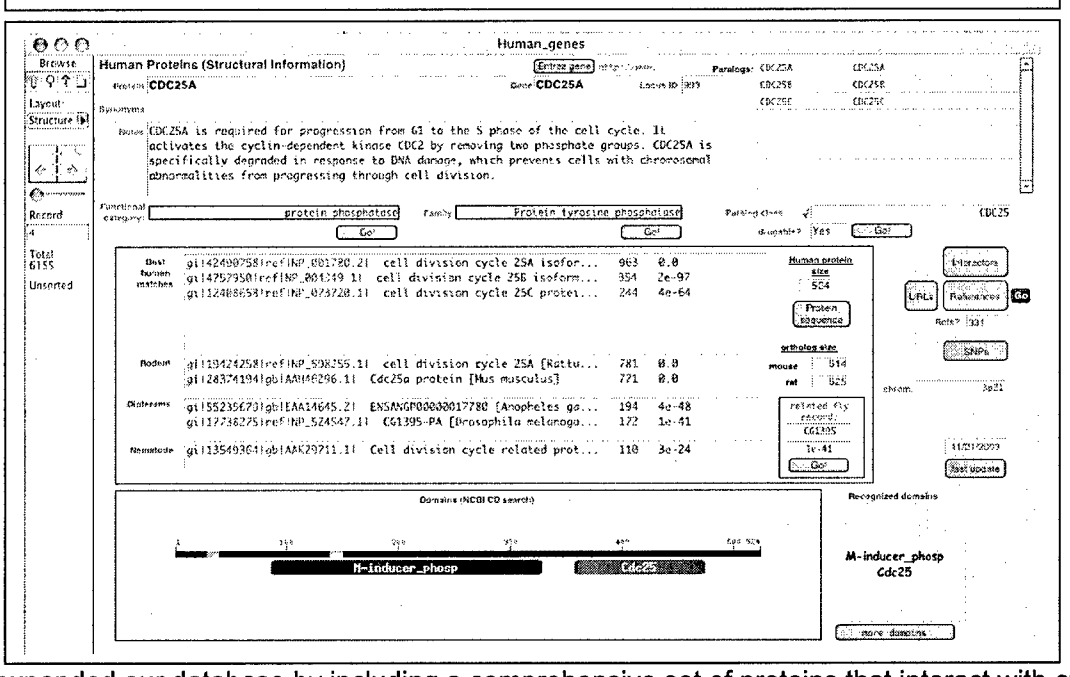


Figure 2. Layout of human database summarizing structural information, in this case for the *CDC25A* protein phosphatase. The portal window (upper right) identifies paralogs.

NF1 modifiers might have roles in signal transduction that are not immediately linked to Ras-like GTPases. For this reason we further expanded our database by including a comprehensive set of proteins that interact with effectors, GAPs, and GEFs, as well as a comprehensive collection of general signaling proteins, such as various receptors,

protein and lipid kinases, phosphatases, adaptors, among multiple other classes of proteins. Finally, we merged all information on all three potential classes of modifiers into a single pair of extensively hyperlinked human/Drosophila relational databases. The creation of these database has involved considerable effort, as is illustrated by the fact that the human database currently includes 6157 manually generated gene records, representing roughly one quarter of all human genes. The comprehensive nature of the database is further illustrated by its current size (509 megabytes), and by the fact that it includes abstracts or other information documenting 6166 unique protein-protein interactions (Figure 1). To provide easy access to other information sources and to allow a check on accuracy, records are hyperlinked to matching records of both Entrez Gene and the Human Protein Reference Database (Peri et al., 2004). Currently, 1401 records include URLs for customized PubMed searches, and 47% of records (2881) include results of BLAST searches, used to identify paralogs and orthologs (Figure 2).
While we have aimed for inclusiveness in creating our database, only about 1000 of the included genes (about 4% of all genes) fall into the categories that we plan to investigate as potential NF1 tumor burden modifiers. While it is impossible to succinctly summarize why each of these genes was flagged as a potential modifier, among them are 319 genes in the genome stability category and 17 genes residing within the NF1 microdeletion. The remainder represent various categories of signaling protein, including approximately 200 GAPs and GEFs for Ras, Rap, Ral, and Rho GTPases, approximately 100 GTPase effector proteins, and multiple proteins involved in signaling pathways that may be regulated by neurofibromin, such as pathways downstream of the EGF receptor, the Kit stem cell factor receptor, and the NGF receptor. In our database the EGF and Kit receptors are documented as interacting with 111 and 37 different proteins, respectively. Based on evidence in the scientific literature each candidate modifier has been assigned a high, medium, or low priority score. Our main goal is to genotype common missense SNPs in all of the approximately 350 high priority genes in the 600 member NF1 patient exploration panel. Only SNPs showing a positive association with high or low tumor burden would also be genotyped in the independent 600 member validation panel.

Progress towards designing and validation of SNP genotyping assays:

Aim 3 for years 1 and 2 was the design and validation of SNP genotyping assays. In our previous report we noted that in order to obtain experience with practical aspects of high throughput genotyping and data analysis, we had sought separate funding to allow us to genotype available case-control panels representing early onset (diagnosis <40 years) breast cancer patients. This work was done to test the hypothesis whether genome stability genes play roles as modifiers of breast cancer susceptibility. Similar to our proposed strategy in the NF1 study, we genotyped separate exploration and validation patient/control panels for common missense SNPs in candidate modifier genes. Initially using a single nucleotide primer extension fluorescence polarization genotyping assay (Kwok, 2002), later superseded by more robust methods, such as direct sequence analysis, allele specific PCR, or RFLP analysis, we performed >20,000 genotyping assays, analyzing missense SNPs in several genes, including *BRCA1, BRCA2, ATM, TP53*, and several Fanconi Anemia (FA) genes. Several observations implicate FA genes as potential breast cancer susceptibility modifiers. Among the most compelling findings, biallelic *BRCA2* mutations have been reported to characterize patients of the *FANCD1* complementation group (Howlett et al., 2002), and *FANCA*, which is the most frequently mutated FA gene, maps to a 650 kb interval which in a recent meta-analysis of breast cancer loss-of-heterozygosity studies was identified as having the second most significant *P* value (Miller et al., 2003). Interestingly, we found that several *FANCA* SNPs showed highly significant differences in genotype distributions between cases and controls in both exploration and validation panels. For example, the two-degree of freedom *P* value for the *FANCA* T266A SNP is 0.02 in both the exploration and validation panels, resulting in a combined P value of 0.002. Another measure of statistical significance, the odds ratio of the cancer associated homozygous T266A genotype is 1.75, with a 95% confidence interval ranging from 1.32 to 2.32. Since several *FANCA* SNPs showed similar associations, we genotyped a total of 16 SNPs in *FANCA* and three immediately adjacent genes and collaborated with statistical geneticist Dr. Mark Daly at MIT to determine pair-wise linkage disequilibrium (LD) patterns and to identify a disease associated SNP haplotype (Figure 3). Importantly, 25% of controls and 38% of patients in both our panels are homozygous for the breast cancer associated SNP haplotype. It is important in this respect that several SNPs in the haplotype are known to have different allele frequencies in different ethnic populations. However, this does not provide an explanation for the large

differences observed between patients and controls, because at least 95% of cases and controls in both our panels are self-reported Caucasians. We are excited about this finding, because most familial clustering of breast cancer remains to be explained (Antoniou et al., 2002), and since our finding indicates that one quarter of Caucasians are homozygous for a *FANCA* SNP haplotype that may confer a 1.75 odds ratio for developing early onset breast cancer.

To confirm the significance of our finding, we performed several functional assays with lymphoblast lines representing breast cancer associated and non-associated SNP haplotypes. No obvious deficiencies were found in experiments that analyzed *FANCA* mRNA levels by real-time PCR, aberrant splicing (one associated SNP maps near a splice branch site), or sensitivity to the DNA cross-linking agent mitomycin-C, which is characteristically elevated in cells from FA patients. The latter assay was done in collaboration with Dr. Alan d'Andrea at HMS. The fact that no functional correlate was found may reflect the fact that lymphoblasts are not breast cancer precursors. However, since we did not find a functional reason for the detected genetic association, we are currently collaborating with Drs. Matthew Freedman and David Althuler (MGH), and with Dr. Laura van het Veer (Netherlands Cancer Institute), to provide further genetic confirmation in additional European and North American breast cancer patients, including patients with different age of disease onset. We have presented these results in some detail, since they document our capacity to do this kind of study and illustrate how we would proceed to follow-up statistically significant allele associations in high and low tumor burden NF1 patients.

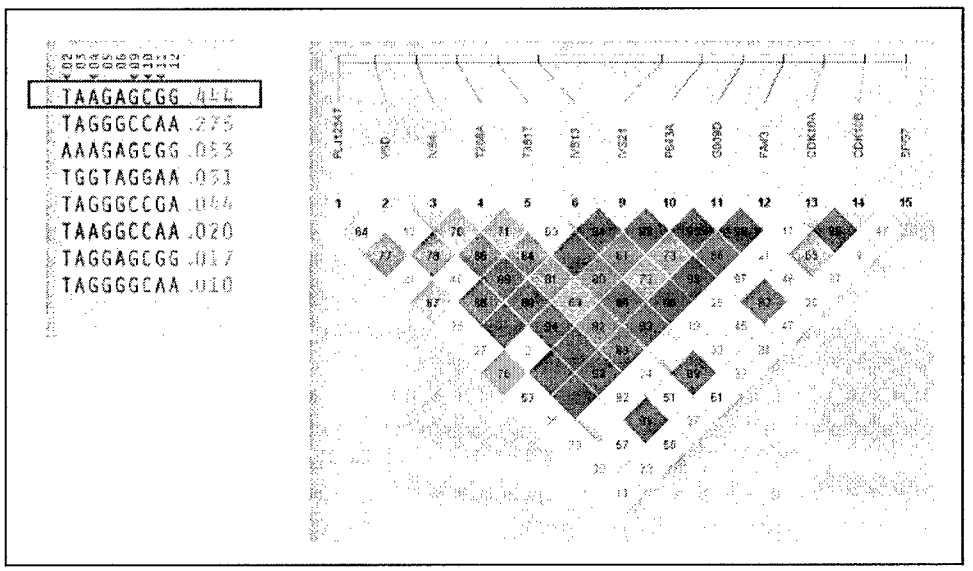


Figure 3. Left panel: Frequency of *FANCA* haplotypes in validation panel. Individuals homozygous for the most common haplotype are over-represented among patients. Right panel: Pair-wise LD patterns across 15 SNPs in *FANCA* and the adjacent *FLJ12547*, *CDK10*, and *SPG7* genes. Red/pink coloration indicates statistical significance (LOD=3). The number in each box is the value of D' * 100. Boxes without a number indicate D'=1 (complete LD). Blue boxes indicate D'=1, but no statistical significance due to low numbers.

Progress on aim 4: Although we have not yet reached the 600 patient recruitment goal that we had identified as the start of the genotyping phase, we have begun genotyping selected high priority SNPs during the past year. While 240 patients is unlikely to allow the detection of statistically significant genotype differences between low and high tumor burden patients, we started genotyping to generate preliminary data and to help convince clinical collaborators of the urgent need for additional patients. SNPs that we have analyzed include two *FANCA* missense polymorphisms (T266A and G501S), the first of which tags the breast cancer associated haplotype. Both SNPs were genotyped by restriction fragment length polymorphism analysis. We also genotyped the recently reported SNP309 in the MDM2 promoter by direct sequence analysis of PCR amplified genomic DNA. SNP309 affects p53-dependent apoptotic signaling by affecting the level of MDM2 expression. We analyzed SNP309 because individuals homozygous for the G allele, which show higher levels of MDM2 expression, have been reported to exhibit accelerated tumor development (ref). So far our results show no obvious difference in FANCA or MDM2 genotypes between high and low neurofibroma burden patients, but the number of patients genotyped is insufficient to make this a firm conclusion.

Progress on aim 5: In our original grant application we proposed to use a single base extension fluorescent polarization genotyping method to analyze samples. Our breast cancer related work has demonstrated that this method is insufficiently robust to allow high throughput genotyping of multiple SNPs. Rather we have genotyped SNPs using labor intensive manual methods, such as RFLP analysis, allele-specific PCR, or direct sequencing. While these methods generate high quality data, they are too labor intensive and too expensive to allow the determination of the required number of genotypes in this project. A attractive current high throughput robotic and relatively low cost (currently $0.17/SNP) genotyping method is offered by the MIT/Broad Institute Genome Center (Oliphant et al., 2002). This Illumina beadarray technology genotyping is especially useful when analyzing large numbers of samples, and we will further explore its use when sufficient patients are recruited.

Key Research Accomplishments:

- Generated an extensively hyperlinked relational database of 6157 human genes.
- Identified and prioritized approximately 1000 candidate neurofibroma burden modifiers.
- Demonstrated proficiency in genotype and haplotype analysis, and in statistical analysis of data.
- Started genotyping high priority SNPs in low and high neurofibroma burden NF1 patients
- Identified Illumina BeadArray Genotyping as method of choice for large-scale genotyping.

Reportable outcomes

Publications:

Bernards A, Settleman J. GAP control: Regulating the regulators of small GTPases. Trends Cell Biol. 2004, 14:377-85.

Bernards A, Settleman J GAPs in growth factor signaling. Growth factors 2005, in press.

Bernards A. Ras Superfamily and interacting proteins database. Methods in Enzymology. Regulators and Effectors of Small GTPases, Part D. Ras Family, C. Der, Ed. 2005, in press.

Databases:

Ras Superfamily GTPase and Affiliated Proteins Database. Filemaker format database described in Methods in Enzymology paper will be made available online.

Abstracts:

A Ras Superfamily GTPase and Affiliated Proteins Database. A. Bernards. Presented at 2004 FASEB meeting in Snowmass, CO.

Identification of Genetic Modifiers of Neurofibroma Burden in NF1. Andre Bernards, Shivang Shah, Qiujuan Wang and Conxi Lazaro. Presented at 2004 meeting of the International Consortium for the Molecular Biology of NF1 and NF2 in Aspen, CO.

Identifying genetic modifiers of neurofibroma development. Andre Bernards and Conxi Lazaro. Presented at the 2005 2004 meeting of the International Consortium for the Molecular Biology of NF1, NF2 and Schwannomatosis in Aspen, CO.

Conclusions

At the start of this project we experienced an unexpected 8 month delay in receiving regulatory approval. Patient recruitment was also adversely impacted by the loss of one major collaborator, and by

the fact that some promised patient samples were not eligible or otherwise not useable. We have taken various measures to make up for the resulting shortfall in patient recruitment and we have enrolled or are in the process of enrolling several new collaborators. In the bioinformatics phase of this study we have generated a relational database of 6157 human genes, and we have used this database to identify approximately 1000 genes (4% of the total gene number) as candidate neurofibroma burden modifiers. We have gained experience with genotyping and data analysis in a spin off project to analyze the role of genome stability genes in breast cancer development, and have started genotyping selected high priority SNPs in high and low neurofibroma burden patients. We have also selected the Illumina platform as the current high throughput genotyping procedure of choice and we plan to use this method to scale-up genotyping when sufficient patients become available.

References:

Antoniou, A. C., Pharoah, P. D., McMullan, G., Day, N. E., Stratton, M. R., Peto, J., Ponder, B. J., and Easton, D. F. (2002). A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer* 86, 76-83.

Arva, N. C., Gopen, T. R., Talbott, K. E., Campbell, L. E., Chicas, A., White, D. E., Bond, G. L., Levine, A. J., and Bargonetti, J. (2005). A chromatin associated and transcriptionally inactive p53-MDM2 complex occurs in MDM2 SNP 309 homozygous cells. *J Biol Chem.*

Bernards, A. (2005). Ras Superfamily and Interacting Proteins Database. In Regulators and Effectors of small GTPases, C. Der, ed. (Amsterdam, Elsevier), pp. in press.

Bernards, A., and Settleman, J. (2004). GAP control: regulating the regulators of small GTPases. *Trends Cell Biol* 14, 377-385.

Bernards, A., and Settleman, J. (2005). GAPs in growth factor signaling. *Growth Factors* In press.

Bond, G. L., Hu, W., Bond, E. E., Robins, H., Lutzker, S. G., Arva, N. C., Bargonetti, J., Bartel, F., Taubert, H., Wuerl, P., et al. (2004). A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* 119, 591-602.

Easton, D. F., Ponder, M. A., Huson, S. M., and Ponder, B. A. (1993). An analysis of variation in expression of neurofibromatosis (NF) type 1 (NF1): evidence for modifying genes. *Am J Hum Genet* 53, 305-313.

Howlett, N. G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G., et al. (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 297, 606-609.

Huson, S. M., and Hughes, R. A. C., eds. (1994). The Neurofibromatoses: A pathogenetic and clinical overview, First edn (London, Chapman & Hall Medical).

Jenne, D. E., Tinschert, S., Reimann, H., Lasinger, W., Thiel, G., Hameister, H., and Kehrer-Sawatzki, H. (2001). Molecular characterization and gene content of breakpoint boundaries in patients with neurofibromatosis type 1 with 17q11.2 microdeletions. *Am J Hum Genet* 69, 516-527.

Kwok, P. Y. (2002). SNP genotyping with fluorescence polarization detection. *Hum Mutat* 19, 315-323.

Miller, B. J., Wang, D., Krahe, R., and Wright, F. A. (2003). Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions. *Am J Hum Genet* 73, 748-767.

Oliphant, A., Barker, D. L., Stuelpnagel, J. R., and Chee, M. S. (2002). BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* Suppl, 56-58, 60-51.

Peri, S., Navarro, J. D., Kristiansen, T. Z., Amanchy, R., Surendranath, V., Muthusamy, B., Gandhi, T. K., Chandrika, K. N., Deshpande, N., Suresh, S., et al. (2004). Human protein reference database as a discovery resource for proteomics. *Nucleic Acids Res* 32, D497-501.

Szudek, J., Birch, P., Riccardi, V. M., Evans, D. G., and Friedman, J. M. (2000). Associations of clinical features in neurofibromatosis 1 (NF1). *Genet Epidemiol* 19, 429-439.

Szudek, J., Evans, D. G., and Friedman, J. M. (2003). Patterns of associations of clinical features in neurofibromatosis 1 (NF1). *Hum Genet* 112, 289-297.

Szudek, J., Joe, H., and Friedman, J. M. (2002). Analysis of intrafamilial phenotypic variation in neurofibromatosis 1 (NF1). *Genet Epidemiol* 23, 150-164.

Ras superfamily and interacting proteins database.

Andre Bernards

Massachusetts General Hospital Center for Cancer Research

Harvard Medical School

149 13th Street, Charlestown, MA 02129. USA

abernards@helix.mgh.harvard.edu

Tel: (617) 726-5620

Facsimile: (617) 724-9648

Running title: Ras superfamily database

Abstract

For geneticists and other researchers alike it is often useful to know how many related proteins may perform similar functions. With this in mind a survey was performed to determine what proportion of human and Drosophila genes code for Ras Superfamily members and their positive or negative regulators. Results indicate that just under 2% of genes in both genomes predict such proteins. A database was compiled to provide easy access to this information. This database also includes information on approximately 360 putative Ras superfamily effector proteins, and may be a useful tool for those interested in GTPase biology.

I. Introduction

Members of the Ras superfamily of small GTP binding proteins control a variety of biological processes by cycling between GDP- and GTP-bound conformational states. The conversion between inactive GDP- and active GTP-bound states is promoted by two main classes of regulatory protein, the Guanine nucleotide Exchange Factors (GEFs), and the GTPase Activating Proteins (GAPs). The former stabilize the nucleotide-free state of Ras superfamily members and promote the exchange of GDP with the more abundant GTP, whereas GAPs stimulate the low intrinsic GTPase activity of many Ras-like proteins, thus causing their inactivation. In their active GTP-bound state, Ras superfamily members can interact with a diverse set of so-called effector proteins, which mediate their various biological responses (Takai et al., 2001).

Members of the Ras family, their regulators and effectors play important roles in many biological processes, and defects in some of the corresponding genes have been implicated in a variety of human diseases, ranging from developmental, neurological, and immunological disorders to inherited and sporadic forms of cancer (Bernards and Settleman, 2004). It is hardly surprising, therefore, that since their discovery 25 years ago, Ras superfamily members and their associated proteins have been the subject of intense scrutiny.

In biological research it is often important to know how many genes or proteins may potentially perform similar functions. While for most of the past 25 years this question could not be answered, the availability of several nearly complete genome sequences presently allows a more definitive accounting of the extent of gene families. Thus, it is now possible to estimate how many Ras superfamily members are encoded by, for example, the human genome (Colicelli, 2004). Since many GEFs and GAPs for the Arf, Rab, Ran, Ras, and Rho branches of the Ras superfamily can be recognized by virtue of their characteristic catalytic domains, it is also possible to estimate what proportion of genomes is devoted to these regulatory proteins. Finally, a still increasing number of GTPase-binding, potential effector proteins continues to be described in the literature. For example, over 60 different proteins, not including GAPs and GEFs, have been reported to interact with Rac1 alone. Therefore, to provide a comprehensive picture of the complexity of biological processes involving Ras superfamily members, it would also be interesting to survey and catalog the universe of potential effector proteins.

II. A database of Ras Superfamily proteins and their regulators.

To identify evolutionary conserved GAPs, I previously explored how many proteins related to Ras superfamily GAPs are encoded by the human and Drosophila genomes. Using a combination of literature and reiterative cross-species BLAST searches, this two year old survey found 173 human and 64 Drosophila GAP-like genes, representing approximately 0.5% of all genes in either species (Bernards, 2003). Importantly, although only about half of the potential human GAPs had been functionally analyzed, at least 85% of the studied proteins were determined to be active GAPs (Bernards and Settleman, 2004).

A similar combination of literature and cross-species BLAST searches has since been used to additionally identify comprehensive sets of Ras superfamily proteins and their GEFs. Results of this expanded survey have again been entered into more elaborate versions of the human and Drosophila databases. In either database each gene-specific record is linked to its closest relative in the other, and records are also hyperlinked to online resources, such as Pubmed, Entrez Gene, Fly Base, and the Human Protein Reference Database. Demonstration versions of the databases are available at http://www.massgeneral.org/cancer/cancer_ccr_Bernards.asp. These versions do not require specialized software, but cannot be modified. Fully functional and modifiable databases that require the Filemaker Pro 7 (Windows or Mac) application are also available upon request.

The most basic function of the human and Drosophila databases is to provide overviews of the GTPase, GAP and GEF gene families. Thus, Figure 1 lists 16 out of 68 identified putative human RhoGAPs grouped by structural similarity. Figure 2 shows an individual record for human RASAL1, a member of the Gap1 family of RasGAPs that was recently found to undergo, and to be regulated by, Ca^{++}-dependent membrane oscillations (Walker et al., 2004). Records can be viewed in several layouts, which provide access to structural information (Figure 2), literature references, or to data about protein interactors (Figure 3, see below).

Since protein structure is an unreliable predictor of function, it is easier to compile lists of structurally related proteins than it is to assign proteins to functional categories. However, among 482 genes in the latest version of the human database, 159 predict Ras superfamily members, 172 predict GAPs and GAP-like proteins, and 155 code for putative or confirmed GEFs (Table 1). The total number of genes is less than the sum of these three categories, since three proteins combine GAP and GEF domains, whereas ARD1 exhibits GAP activity towards its own Arf-like GTPase domain. Nevertheless, the fact that roughly 2% of the estimated 25,000 human genes predict proteins related to Ras superfamily members and their GAP or GEF regulators, serves to reemphasize the critical importance of the Ras superfamily.

The numbers in Table 1 are best estimates and may change for several reasons. Thus, some identified genes may in fact be pseudogenes, and some excluded pseudogenes may turn out to be functional. More importantly, yet to be discovered

regulators may be unrelated to presently known GAPs or GEFs. This point is illustrated by the recent discovery that members of the Dock family, which lack obvious similarity to previously characterized GEFS, can serve as GEFs for Rac, CDC42, and perhaps Rap GTPases (Brugnera et al., 2002; Cote and Vuori, 2002; Meller et al., 2002; Namekata et al., 2004; Nishikimi et al., 2005; Yajnik et al., 2003). Moreover, although distant similarity between the cytoplasmic segments of plexins and the catalytic domains of RasGAPs had been noted, it was unexpected when plexin-B1 was recently identified as a GAP for R-Ras (Oinuma et al., 2004).

Table 1 includes 9 human plexins as potential RasGAPs and 11 Dock family members as putative RhoGEFs. Also included as potential GEFs are 14 RCC1 repeat proteins. The prototype member of this group, regulator of chromosome condensation 1, functions as the single known GEF for the Ran GTPase (Bischoff and Ponstingl, 1991). Among the 13 other RCC1 repeat proteins, the giant protein HERC1 has been reported to serve as a GEF for Arf1, Arf6, and Rab2 (Rosa et al., 1996), whereas the chromosomal passenger protein TD-60 was recently found to interact with the nucleotide-free form of Rac1 (Mollinari et al., 2003).

Beyond providing an overview of Ras superfamily members and their main regulators, the human database has been designed to serve as a repository of information reported in the scientific literature. Thus, each record includes a script to run a tailored PubMed search for the gene/protein in question, and records allow storage of any relevant information. For example, a layout specific to GAPs or GEFs can store information on up

to 20 potential GTPase substrates. Database records for Ras superfamily members query this information and use it to identify specific GAPs and GEFs. In addition, each GTPase, GAP or GEF record can also display information on up to 60 potential interacting proteins. For example, Figure 3 identifies 7 GAPs and 9 GEFs that have been reported to display activity towards H-Ras, in addition to 32 proteins that interact with this GTPase. Literature references documenting these interactions are accessed by clicking the name of the GAP or GEF, or the numbered buttons to the right of individual interactors.

The total number of human Ras superfamily interactors entered into the database is just under 360. This number largely represents proteins that physically interact with Ras superfamily members, although a small number of functional interactors is also included. It is important to note that this information is unlikely to be complete or entirely accurate. Thus, apart from the challenge of summarizing sometimes contradictory information on 2-3% of human genes, authors often do not specify exactly which paralog has been analyzed. In such cases we have assumed that the most commonly used paralog, for example RhoA or H-Ras, was being analyzed, although this may not always have been the case. To provide a check on accuracy, database records have been linked to matching records in the online Human Protein Reference Database, which also includes information on interacting proteins (Peri et al., 2004). However, until feedback from other researchers has been incorporated, any specific information should be used with caution.

Literature cited:

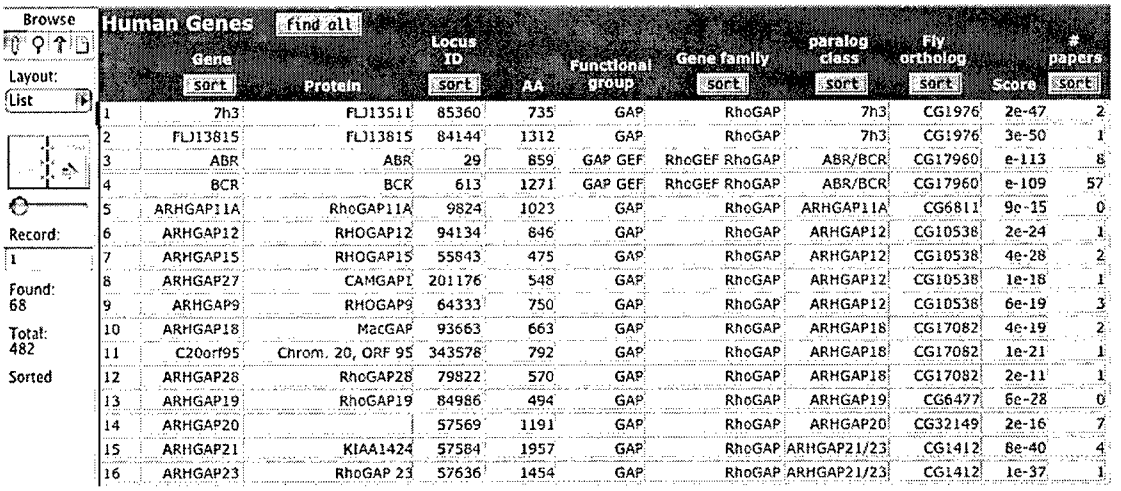
- Bernards, A. (2003). GAPs galore! A survey of putative Ras superfamily GTPase activating proteins in man and Drosophila. Biochim Biophys Acta *1603*, 47-82.
- Bernards, A., and Settleman, J. (2004). GAP control: regulating the regulators of small GTPases. Trends Cell Biol *14*, 377-385.
- Bischoff, F. R., and Ponstingl, H. (1991). Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature *354*, 80-82.
- Brugnera, E., Haney, L., Grimsley, C., Lu, M., Walk, S. F., Tosello-Trampont, A. C., Macara, I. G., Madhani, H., Fink, G. R., and Ravichandran, K. S. (2002). Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex. Nat Cell Biol *4*, 574-582.
- Colicelli, J. (2004). Human RAS superfamily proteins and related GTPases. Sci STKE *2004*, RE13.
- Cote, J. F., and Vuori, K. (2002). Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J Cell Sci *115*, 4901-4913.
- Meller, N., Irani-Tehrani, M., Kiosses, W. B., Del Pozo, M. A., and Schwartz, M. A. (2002). Zizimin1, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. Nat Cell Biol *4*, 639-647.
- Mollinari, C., Reynaud, C., Martineau-Thuillier, S., Monier, S., Kieffer, S., Garin, J., Andreassen, P. R., Boulet, A., Goud, B., Kleman, J. P., and Margolis, R. L. (2003). The mammalian passenger protein TD-60 is an RCC1 family member with an essential role in prometaphase to metaphase progression. Dev Cell *5*, 295-307.

- Namekata, K., Enokido, Y., Iwasawa, K., and Kimura, H. (2004). MOCA induces membrane spreading by activating Rac1. J Biol Chem *279*, 14331-14337.
- Nishikimi, A., Meller, N., Uekawa, N., Isobe, K., Schwartz, M. A., and Maruyama, M. (2005). Zizimin2: a novel, DOCK180-related Cdc42 guanine nucleotide exchange factor expressed predominantly in lymphocytes. FEBS Lett *579*, 1039-1046.
- Oinuma, I., Ishikawa, Y., Katoh, H., and Negishi, M. (2004). The Semaphorin 4D receptor Plexin-B1 is a GTPase activating protein for R-Ras. Science *305*, 862-865.
- Paulding, C. A., Ruvolo, M., and Haber, D. A. (2003). The Tre2 (USP6) oncogene is a hominoid-specific gene. Proc Natl Acad Sci U S A *100*, 2507-2511.
- Peri, S., Navarro, J. D., Kristiansen, T. Z., Amanchy, R., Surendranath, V., Muthusamy, B., Gandhi, T. K., Chandrika, K. N., Deshpande, N., Suresh, S., *et al.* (2004). Human protein reference database as a discovery resource for proteomics. Nucleic Acids Res *32*, D497-501.
- Rosa, J. L., Casaroli-Marano, R. P., Buckler, A. J., Vilaro, S., and Barbacid, M. (1996). p619, a giant protein related to the chromosome condensation regulator RCC1, stimulates guanine nucleotide exchange on ARF1 and Rab proteins. Embo J *15*, 4262-4273.
- Takai, Y., Sasaki, T., and Matozaki, T. (2001). Small GTP-binding proteins. Physiol Rev *81*, 153-208.
- Walker, S. A., Kupzig, S., Bouyoucef, D., Davies, L. C., Tsuboi, T., Bivona, T. G., Cozier, G. E., Lockyer, P. J., Buckler, A., Rutter, G. A., *et al.* (2004). Identification of a Ras GTPase-activating protein regulated by receptor-mediated Ca(2+) oscillations. Embo J *23*, 1749-1760.

• Yajnik, V., Paulding, C., Sordella, R., McClatchey, A. I., Saito, M., Wahrer, D. C., Reynolds, P., Bell, D. W., Lake, R., van den Heuvel, S., *et al.* (2003). DOCK4, a GTPase activator, is disrupted during tumorigenesis. Cell *112*, 673-684.

GTPase family	Ras Superfamily members	GAPs	GEFs
Total	159	172	155
Arf	31	ArfGAP domain: 24 Other: 1	Sec7 domain: 16
Rab	69	TBC domain: 44 Other: 1	VPS9 domain: 8 Other: 4
Ran	1	1	1
Ras	37	RasGAP domain: 24 RapGAP domain: 10	Cdc25 domain: 32
Rho	19	RhoGAP domain: 67 Other: 1	Dbl domain: 72 Dock-related: 11
Sar	2	SEC23-like: 2	SEC12-like: 1
Miscellaneous			RCC1-repeat: 13 Smg-GDS: 1

Table 1. Number of human genes in the indicated categories. These numbers are best estimates, and should not be seen as definitive. For complete information and references see the human database. Three GAPs contain both ArfGAP and RhoGAP domains. Four GEFs include both Cdc25 and Dbl domains. Several nearly identical copies of the *CTGLF1* gene, predicting a centaurin-gamma-like putative ArfGAP, exist in the human genome, but only *CTGLF1* itself is included in the above numbers. Several highly related paralogs of the *TBC1D3* gene, encoding a putative human RabGAP, have been included (Paulding et al., 2003).



Browse

Layout: List

Record: 1

Found: 68

Total: 482

Sorted

Human Genes find all

	Gene sort	Protein	Locus ID sort	AA	Functional group	Gene family sort	paralog class sort	Fly ortholog sort	Score	# papers sort
1	7h3	FLJ13511	85360	735	GAP	RhoGAP	7h3	CG1976	2e-47	2
2	FLJ13815	FLJ13815	84144	1312	GAP	RhoGAP	7h3	CG1976	3e-50	1
3	ABR	ABR	29	859	GAP GEF	RhoGEF RhoGAP	ABR/BCR	CG17960	e-113	8
4	BCR	BCR	613	1271	GAP GEF	RhoGEF RhoGAP	ABR/BCR	CG17960	e-109	57
5	ARHGAP11A	RhoGAP11A	9824	1023	GAP	RhoGAP	ARHGAP11A	CG6811	9e-15	0
6	ARHGAP12	RHOGAP12	94134	846	GAP	RhoGAP	ARHGAP12	CG10538	2e-24	1
7	ARHGAP15	RHOGAP15	55843	475	GAP	RhoGAP	ARHGAP12	CG10538	4e-28	2
8	ARHGAP27	CAMGAP1	201176	548	GAP	RhoGAP	ARHGAP12	CG10538	1e-18	1
9	ARHGAP9	RHOGAP9	64333	750	GAP	RhoGAP	ARHGAP12	CG10538	6e-19	3
10	ARHGAP18	MacGAP	93663	663	GAP	RhoGAP	ARHGAP18	CG17082	4e-19	2
11	C20orf95	Chrom. 20, ORF 95	343578	792	GAP	RhoGAP	ARHGAP18	CG17082	1e-21	1
12	ARHGAP28	RhoGAP28	79822	570	GAP	RhoGAP	ARHGAP18	CG17082	2e-11	1
13	ARHGAP19	RhoGAP19	84986	494	GAP	RhoGAP	ARHGAP19	CG6477	6e-28	0
14	ARHGAP20		57569	1191	GAP	RhoGAP	ARHGAP20	CG32149	2e-16	7
15	ARHGAP21	KIAA1424	57584	1957	GAP	RhoGAP	ARHGAP21/23	CG1412	8e-40	4
16	ARHGAP23	RhoGAP 23	57636	1454	GAP	RhoGAP	ARHGAP21/23	CG1412	1e-37	1

Figure 1. List view of human database, showing 16 of 68 human RhoGAPs sorted by structural similarity.

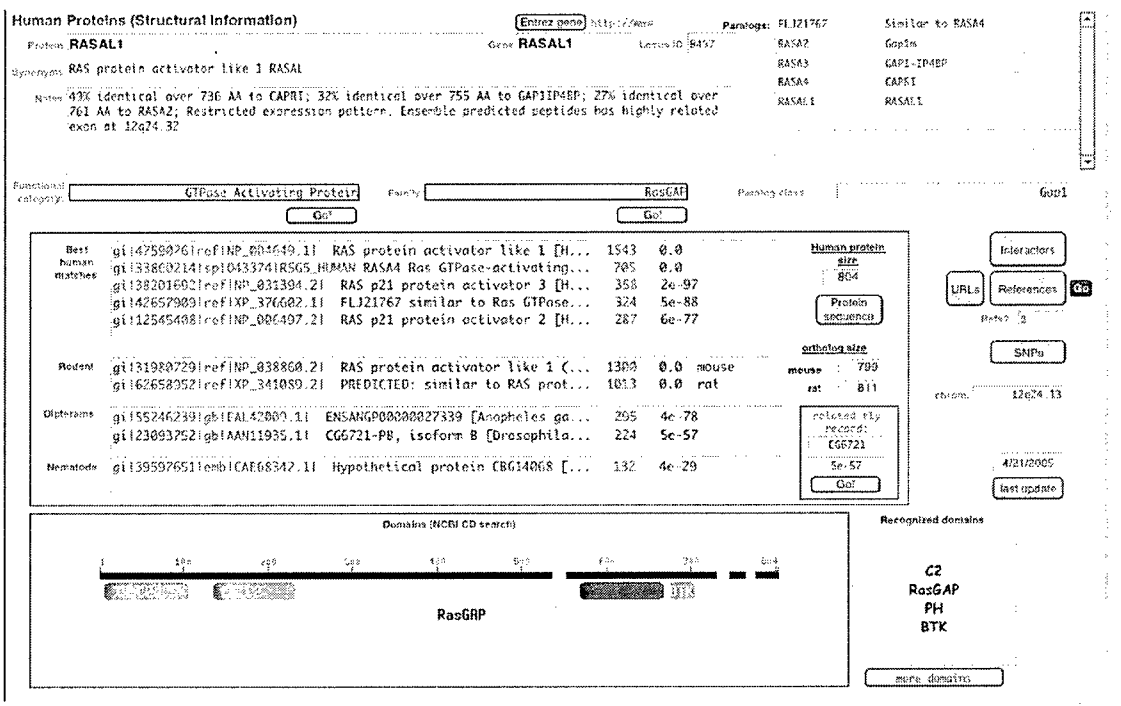


Figure 2. Individual record view of human database, showing structural information for human RASAL1. The portal window in the upper right hand corner identifies and provides access to related human genes.

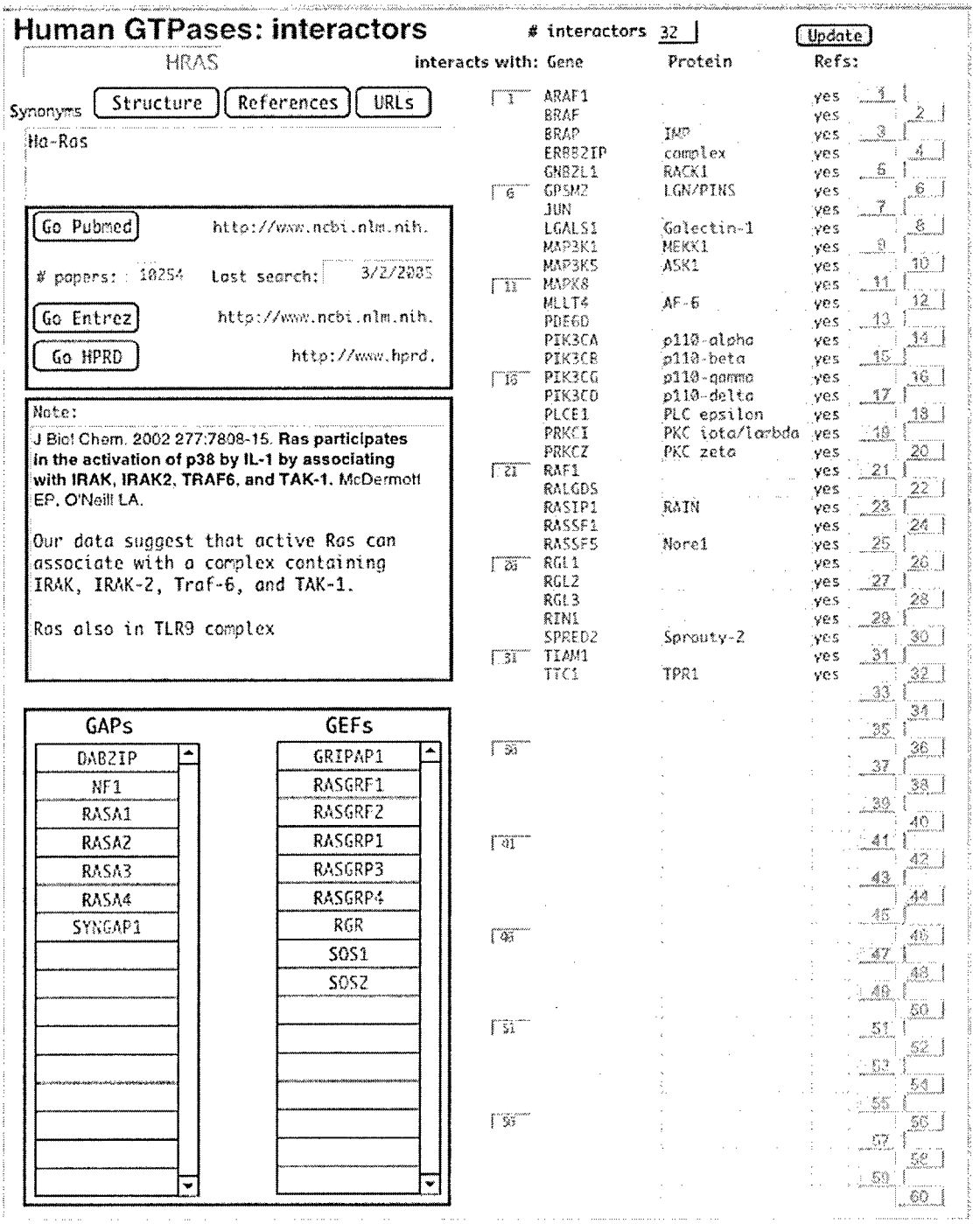


Figure 3. Human database layout that summarizes information about GTPase interactors. Seven GAPs, nine GEFs, and 32 putative effectors and other proteins have been reported to interact with H-Ras.

Review TRENDS in Cell Biology Vol.not known No.not known Month 0000

GAP control: regulating the regulators of small GTPases

Andre Bernards and Jeffrey Settleman

MGH Cancer Center and Harvard Medical School, 149 13th Street, Charlestown, MA 02129, USA

The small GTPases of the Ras superfamily mediate numerous biological processes through their ability to cycle between an inactive GDP-bound and an active GTP-bound form. Among the key regulators of GTPase cycling are the GTPase-activating proteins (GAPs), which stimulate the weak intrinsic GTP-hydrolysis activity of the GTPases, thereby inactivating them. Despite the abundance of GAPs and the fact that mutations in GAP-encoding genes underlie several human diseases, these proteins have received relatively little attention. Recent studies have addressed the regulatory mechanisms that influence GAP activity. So far, findings suggest that GAP activity is regulated by several mechanisms, including protein–protein interactions, phospholipid interactions, phosphorylation, subcellular translocation and proteolytic degradation.

There are at least 140 small GTPases encoded by the human genome, and the various subclasses of this protein superfamily (including the Ras, Rho, Arf, Rab and Ran GTPases) have been implicated in almost all aspects of cell biology, including proliferation, differentiation, cytoskeletal organization, vesicle trafficking, nucleocytoplasmic transport and gene expression [1]. These small GTPases can be considered as 'molecular switches', whose cycling between active and inactive forms is regulated stringently by cellular factors [1] (Figure 1).

The guanine nucleotide exchange factors (GEFs) comprise a large family of GTPase regulators that promote the formation of active GTP-bound GTPases, whereas the GAPs promote GTPase inactivation by stimulating GTP-hydrolysis activity. In addition, a class of regulatory proteins, known as guanine nucleotide dissociation inhibitors (GDIs), has been identified for some subclasses of small GTPases. The GDIs seem to function as inhibitors of GTPase activation by preventing the dissociation of GDP from the inactive GTPase.

The current dogma, at least in the context of signal transduction, proposes that GEFs are activated in response to various stimuli and consequently promote GTPase activation. The activated GTPase undergoes a conformational change that enables it to interact with so-called downstream 'GTPASE EFFECTORS', which ultimately produce a biological response. The cycle is then completed through the action of GAPs [2]. The precise role of GDIs in the cycle is less clear.

Corresponding author: Jeffrey Settleman (settleman@helix.mgh.harvard.edu).

This widely held view essentially relegates the GAPs to a 'secondary' role that is seemingly less significant than that of the GEFs, which have received considerably more attention in the context of GTPase regulation. Some studies have indicated, however, that GAPs can potentially function as effectors of activated GTPases [3]. Moreover, inhibition of RhoGAP activity has been found to be sufficient to promote activation of Rho, and a consequent Rho-mediated biological response, in the absence of extracellular stimuli [4], suggesting that GAPs might contribute more to the function of GTPases than is superficially predicted by their biochemical activity. An important biological role for the GAPs is also indicated by the fact that several of them have been implicated in human diseases (Box 1).

The impressively large number of predicted GAPs is consistent with these proteins having a crucial role in GTPase regulation, and their structural features indicate that their activity is likely to be under stringent regulatory control. As we describe below, recent studies of the pathways that influence GAP activity suggest that the GAPs are regulated by numerous mechanisms, including protein–protein interactions, phospholipid interactions, phosphorylation, subcellular translocation and proteolytic degradation.

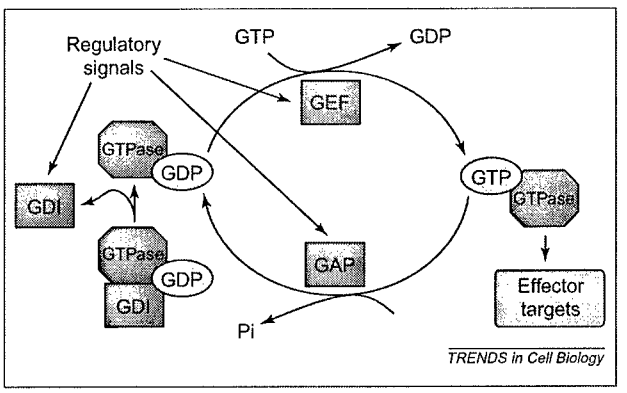


Figure 1. The GTPase cycle. Most small GTPases seem to be regulated in a similar manner as they cycle between their inactive GTP-bound and active GTP-bound forms. The activated GTPase can interact with so-called 'effector targets' that ultimately produce a biological consequence. Regulation of cycling is largely accomplished through the coordinated action of GTPase-activating proteins (GAPs), guanine nucleotide exchange factors (GEFs) and guanine nucleotide dissociation inhibitors (GDIs), the activity of each of which is potentially modulated in response to various signals. Inactive GTPase or GTPase inactivators are colored red, active GTPase or GTPase activators are colored green.

www.sciencedirect.com 0962-8924/$ - see front matter doi:10.1016/j.tcb.2004.05.003

Glossary

GTPase effectors: a group of proteins that typically bind specifically to the GTP-bound, activated form of the various small GTPases. GTPase effectors are frequently involved in eliciting a biological response to GTPase activation. Many of them are protein kinases, and it seems that their interaction with their active GTPase can lead to both subcellular redistribution and catalytic activation of the kinase.

Phospholipids: polar molecules consisting of a long hydrophobic carbon chain and a hydrophilic phosphate 'head' group. Phospholipids are prominent components of cellular membranes, and several types have been implicated in various cellular processes through their ability to regulate protein function through direct binding.

Ras GTPase superfamily: a large group of GTP-binding proteins of roughly 20–30 kDa that are well conserved in evolution. All Ras GTPases show sequence similarity to the Ras oncoprotein and share the property that they bind GTP and have a relatively weak GTP hydrolyzing activity. In humans, there are at least 140 predicted GTPases in this family, which can be further subdivided on the basis of sequence similarity into several classes including the Ras, Rab, Ran, Rho and Arf GTPases. Note that the various GTPases, GAPs and GAP regulatory proteins, many of which are identified only by name in the text, are clearly defined in the website (http://cancer.mgh.harvard.edu/GAP_review.htm) with detailed descriptions of each protein.

Ubiquitin–proteasome system: a pathway for the targeted degradation of proteins, which represents an important mechanism for regulating protein function. Proteins subject to the ubiquitin–proteasome system are typically targeted by the covalent attachment of a polyubiquitin chain, which is then recognized by the proteasome, a large complex of proteins that carries out proteolytic destruction of the ubiquitinated substrate.

The 'GAP-ome'

A recent survey found that at least 160 human genes are predicted to encode proteins that resemble GAPs for various members of the RAS GTPASE SUPERFAMILY [5]. The fact that about 0.5% of all predicted human genes encode likely GAPs (Table 1) suggests that these proteins have widespread and important roles in GTPase regulation. In fact, there could be many more GAPs encoded by the genome than are predicted by sequence homology searches, because GAPs for many of the small GTPases have not been biochemically established and it remains possible that such GAPs are encoded by distinct sequences that have yet to be identified.

The domain organization of proteins can provide important clues about potential regulatory mechanisms, and many predicted GAPs contain various structural motifs. As shown in Tables 1 and 2, most potential protein-binding and lipid-binding motifs are restricted to one or two classes of GAP, with, for example, protein-binding ankyrin repeats only found in ArfGAPs and diacylglycerol-binding C1 domains restricted to RhoGAPs. The most common motifs among GAPs are Src homology domain 3 (SH3) domains that bind proline-containing peptides, and phosphoinositide-binding pleckstrin homology (PH) domains. Some GAPs also contain potential or confirmed enzymatic domains (Table 3), indicating that these proteins might have additional functions, including effector roles downstream of activated GTPases.

A comparison of human and *Drosophila* GAPs previously showed that the domain organization of most GAPs is evolutionarily conserved [5]. Although the existence of a conserved domain organization can be used to make general predictions about potential regulatory mechanisms, such information reveals little about the intricacies of the regulation of individual GAP proteins. In addition, several functional protein motifs and important regulatory phosphorylation sites cannot be reliably predicted by computer searches and could also contribute to the regulation of GAP activity. Thus, the subtleties of GAP regulation have been largely determined through the biochemical analysis of individual proteins. Below, we describe recent examples of such studies that have begun to elucidate the nature of GAP regulation.

Box 1. GAPs and disease

Two of the most common human genetic disorders associated with an increase in cancer risk, neurofibromatosis type 1 (NF1) and tuberous sclerosis complex (TSC), are caused by mutations that disrupt the function of GTPase-activating proteins (GAPs). The *NF1* gene encodes neurofibromin, which functions as a GAP for Ras and its immediate relatives [53], whereas mutations in *TSC1* or *TSC2* affect the function of a two-subunit GAP for the GTPase Rheb, which functions in the target of rapamycin (TOR) signaling pathway [10].

Mutations in the gene *RASA1*, which encodes the first GAP to be identified, p120 RasGAP, have been recently found in individuals affected with a capillary and arteriovenous malformation syndrome [54]. Loss-of-function mutations in the murine ortholog of p120 RasGAP also cause vascularization defects [55].

Oculocerebrorenal syndrome of Lowe (OCRL1) is an uncommon X-linked disease characterized by mental retardation, congenital cataracts and renal Fanconi syndrome. The OCRL1 protein is a phosphatidylinositol (4,5)-bisphosphate 5-phosphatase that contains a functional RhoGAP domain [56]. Through their ability to control F-actin dynamics, Rho GTPases have an important role in axonal pathfinding and other neuronal functions, which might explain why mutations in the RhoGAP oligophrenin [57], the RhoGEF ARHGEF6 [58], and the Rho effector protein PAK3 [59] are each associated with nonspecific X-linked mental retardation.

Three other RhoGAPs have been implicated in disease, although in each case their causative role requires further confirmation. In brief, the gene *GRAF*, which encodes an oligophrenin-like RhoGAP, has been found to show biallelic mutations in individuals affected with 5q-minus myeloid leukemia [60], the gene *deleted in liver cancer 1(DLC1)* is frequently deleted or epigenetically silenced in hepatocellular carcinoma [61], and the X-linked gene *ARHGAP6* maps to the minimal genomic segment implicated in microphthalmia with linear skin defects [62].

Regulation by protein interaction

Among the best-studied examples of GAPs controlled by protein interaction is the budding yeast two-component GAP Bfa1p–Bub2p for the Rab-like GTPase Tem1p, which activates a signaling pathway that controls mitotic exit [6]. The activity of Bfa1p–Bub2p is regulated in several ways, one of which involves phosphorylation of Bfa1p by the Polo-like kinase Cdc5p [7]. Cdc5p phosphorylation of Bfa1p inhibits the GAP activity of Bfa1p–Bub2p without affecting the binding of Tem1p [8]. By contrast, the Cdc42p effector protein Gic1p interacts with Bub2p and prevents its interaction with Tem1p [9].

These interactions might help to explain the role of Cdc42p and its regulators in mitotic exit. Yeast Bub2p contains a TBC1 RabGAP catalytic domain. Most of the 40 human proteins that are currently predicted to include TBC1 domains have not been studied; thus, it remains unclear what proportion of these proteins function as GAPs, and whether they typically function alone or with partners.

Tuberin, which is encoded by the gene *tuberous sclerosis complex-2* (*TSC2*), is a RapGAP-related protein that functions as a GAP for the Rheb GTPase in the

Table 1. Distribution of putative protein interaction domains in human GAPs

Domain[a]	InterPro ID	Hypothetical function	No. of ArfGAPs 26 (15)[b]	No. of RabGAPs 41 (5)[b]	No. of RapGAPs 10 (4)	No. of RasGAPs 15 (8)[c]	No. of RhoGAPs 68 (35)[d]
Ankyrin	IPR002110	Protein interaction	12 (8)	0	0	0	0
CH	IPR001715	F-actin binding	0	0	0	3 (0)	0
CNH	IPR001180	Regulatory domain/unknown	0	0	1 (0)	0	0
DEP	IPR000591	Unknown	0	0	0	0	2 (0)
FCH	IPR001060	Unknown	0	0	0	0	7 (5)
FF	IPR002713	Protein interaction	0	0	0	0	2 (2)
GIT	IPR006557	Protein interaction (?)	2 (2)	0	0	0	0
GoLoco	IPR003109	Binds Gα proteins	0	0	1 (1)	0	0
IQ	IPR000048	Calmodulin/myosin light chain binding	0	0	0	3 (0)	2 (2)
MyTH4	IPR000857	Present in several myosins	0	0	0	0	1 (0)
PDZ	IPR001478	(C-terminal) peptide binding	0	0	4 (2)	0	2 (0)
PTB	IPR000050	Phosphotyrosine binding	0	4 (2)	0	0	0
RA	IPR000159	Ras/Rap association domain	3 (2)	0	0	0	6 (4)
RUN	IPR004012	Present in Rab/Rap regulators/effectors	0	3 (0)	0	0	0
SAM	IPR001660	Protein interaction domain	2 (2)	0	0	0	4 (4)
SH2	IPR000980	Phosphotyrosine binding	0	0	0	1 (1)	4 (2)
SH3	IPR001452	Binds proline/hydrophobic peptides	2 (2)	1 (0)	0	1 (1)	14 (9)
WW	IPR001202	Proline-rich and/or phospho-Ser/phospho-Thr binding	0	0	0	3 (0)	4 (1)

[a]Domains are identified by name and by their InterPro database identifier (http://www.ebi.ac.uk/interpro/). RanGAP1 and two orthologs of the yeast Sec23 GAP contain no obvious domains and are not included in this or the subsequent tables. Among 160 unique GAPs (for details, see the Human and *Drosophila* GAP Databases: http://cancer.mgh.harvard.edu/GAP_review.htm), 80 have been analyzed at the protein level, 76 are predicted by cDNAs, and 4 are based on computer predictions. Of the 80 proteins analyzed, 67 (84%) show GAP activity towards at least one GTPase. Numbers indicate how many putative GAPs include the indicated domains, with numbers in parentheses referring to proteins with confirmed GAP activity.

[b]The sequence of the human genome predicts more than five highly related paralogs of the pleckstrin homology (PH)-domain-containing and ankyrin-repeat-containing putative ArfGAP MRIP2, and similar numbers of the TBC1D3-like RabGAP and RASA4-like RasGAP paralogs. There are also several genes predicting short protein segments related to the catalytic RhoGAP domain of the breakpoint cluster region (BCR) protein. Because it remains unclear how many of these paralogs represent functional genes, they are not included among the 160 proteins surveyed in this or the subsequent tables. Numbers in the RabGAP column refer to proteins that contain a putative RabGAP TBC1 catalytic domain. Also included is Rab3GAP, which does not contain a TBC1 domain.

[c]Included among RasGAP-like proteins is one member of the plexin family, to make note of the fact that some plexins include cytoplasmic segments related to RasGAP catalytic domains.

[d]Numbers in the RhoGAP column include BNIP2, which predicts an unconventional RhoGAP. Three related human proteins contain both ArfGAP and RhoGAP domains.

Table 2. Distribution of putative lipid-binding domains in human GAPs[a]

Domain	InterPro ID	Hypothetical function	No. of ArfGAPs	No. of RabGAPs	No. of RapGAPs	No. of RasGAPs	No. of RhoGAPs
BAR	IPR004148	Membrane curvature sensor	6 (4)	0	0	0	6 (6)
BTK	IPR001562	PH domain extension	0	0	0	4 (3)	0
C1	IPR002219	Diacylglycerol binding	0	0	0	0	8 (7)
C2	IPR000008	(Ca^{2+}-dependent) phospholipid binding	0	0	0	9 (7)	2 (2)
GRAM	IPR004182	Phosphoinositide binding	0	4 (0)	0	0	0
PH	IPR001849	Phosphoinositide and peptide binding	16 (10)	2 (0)	0	9 (7)	21 (10)
PX	IPR001683	Phosphoinositide binding	0	0	0	0	1 (1)
SEC14	IPR000051	Lipid binding	0	0	0	1 (1)	3 (3)
START	IPR002913	Lipid binding	0	0	0	0	3 (2)

[a]For details, see footnotes to Table 1.

Table 3. Presence of enzymatic and miscellaneous domains in predicted human GAPs[a]

Domain	InterPro ID	Hypothetical function	No. of ArfGAPs	No. of RabGAPs	No. of RapGAPs	No. of RasGAPs	No. of RhoGAPs
GTPase	IPR003575	GTP binding	4 (4)	0	0	0	2 (2)
IPPc	IPR000300	Found in inositol monophosphatases	0	0	0	0	2 (1)
Kinase	IPR000719	Protein kinase domain	1 (0)	0	0	0	0
RasGAPc	IPR000593	Present in IQGAP-like proteins	0	0	0	3 (0)	0
Rhodanese	IPR001763	Found in phosphatases and other proteins	2 (0)	0	0	0	0
RhoGEF	IPR000219	Rho exchange factor catalytic domain	0	0	0	0	2 (2)
TLDc	IPR006571	Found in TBC/LysM domain proteins	0	1 (0)	0	0	0
UCH-2	IPR001394	Present in hominoid-specific fusion gene	0	1	0	0	0
VPS9	IPR003123	Rab exchange factor domain	0	0	0	1 (0)	0

[a]For details, see footnotes to Table 1.

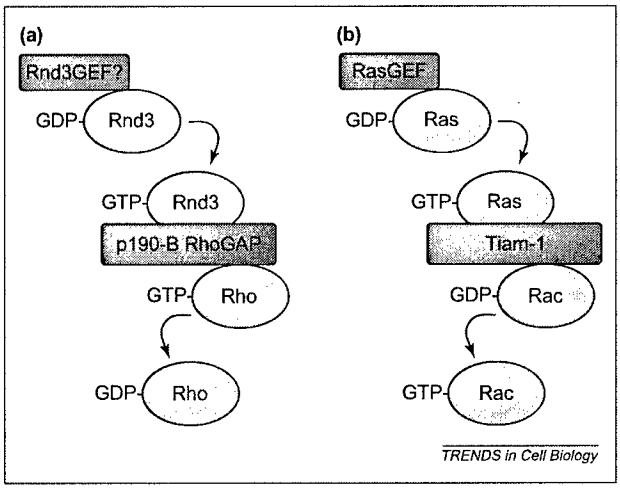


Figure 2. GTPases seem to function in cascades. **(a)** When activated, the GTPase Rnd3 binds directly to p190-B RhoGAP, thereby stimulating its RhoGAP catalytic activity, which in turn promotes inactivation of Rho. **(b)** When activated, the Ras GTPase binds directly to the RacGEF, Tiam-1, thereby promoting its catalytic activity, which in turn promotes the activation of Rac.

phosphatidylinositol 3-kinase (PI3K), Akt, and target of rapamycin (TOR) signaling pathway [10]. Tuberin functions in a complex with hamartin, the protein product of *TSC1*. Complex formation is promoted by phosphorylation [11], which stabilizes tuberin by preventing its ubiquitination [12]. Thus, the essential role of hamartin in the TOR signaling pathway might reflect its ability to prevent the proteasomal degradation of tuberin [10].

Other GAPs that have been suggested to be regulated by protein interactions include RapGAP1, an isoform of which binds several heterotrimeric G-protein α-subunits through an amino (N)-terminal GoLoco motif [13–15]; MgcRacGAP, whose activity is inhibited by the binding of protein regulating cytokinesis 1 (PRC1) to its catalytic domain [16]; and CdGAP, whose RhoGAP activity is inhibited by the binding of the endocytic scaffolding protein intersectin, a brain-specific form of which contains a functional RhoGEF domain [17].

The GAP activity of the p190-B RhoGAP is stimulated by direct interaction with the small GTPase Rnd3 [18], suggesting the existence of 'GTPase cascades' that involve the regulation of one GTPase by another through a GAP mediator (Figure 2). The GEFs have been also implicated in such cascades. For example, the RacGEF Tiam1 can be activated by direct interaction with the activated Ras protein [19] (Figure 2).

Some GAPs are also regulated by intramolecular interactions. For example, the PH domain of p120 RasGAP associates with and regulates the activity of its catalytic domain [20]; and an N-terminal domain of the RhoGAP, oligophrenin, has a role in inhibiting its RhoGAP activity [21]. Many protein–protein interactions are regulated by specific phosphorylations but, as we describe in the next section, phosphorylation of GAPs can also regulate GAP function in other ways.

Regulation by phosphorylation

Protein phosphorylation obviously has a prominent role in the modulation of many proteins, including GAPs. Numerous examples of tyrosine and serine/threonine phosphorylation have been documented and, in several cases, the regulatory effects of such phosphorylation have been determined. GAP phosphorylation has the potential to influence GAP enzymatic function directly through conformational effects on the catalytic site, and it can also affect GAP activity indirectly by regulating the subcellular localization, the targeted degradation and, as described above, the protein interactions of the GAP. Indeed, each of these phosphorylation-mediated regulatory mechanisms has been reported.

Phosphorylation-mediated GAP regulation

Various studies have implicated phosphorylation in the regulation of GAPs, although the precise mechanism remains poorly understood in most cases. For example, the RasGAP neurofibromin (the product of the *NF1* tumor suppressor gene) is phosphorylated at several sites in its carboxy (C)-terminal region by protein kinase A; this phosphorylation promotes the interaction of neurofibromin with 14–3-3 proteins and correlates with a reduction in RasGAP activity [22]. Likewise, the *in vitro* activity of RICS, a Rac and Cdc42 GAP, is inhibited after its phosphorylation by Ca^{2+}/calmodulin-dependent protein kinase II [23]. Notably, this same kinase phosphorylates SynGAP, a neuron-specific RasGAP, causing a modest increase in its catalytic GAP activity [24]. Similarly, insulin-induced Akt phosphorylation of the TBC1-domain-containing protein AS160 has been suggested to inhibit its RabGAP activity and to result in Rab-mediated translocation of the GLUT4 glucose transporter to the plasma membrane [25].

A particularly interesting example of phosphorylation-mediated GAP regulation has been recently reported. MgcRacGAP seems to be involved in cytokinesis and functions as a GAP for Rac and Cdc42, but not Rho. In addition, MgcRacGAP has been found to acquire RhoGAP activity after phosphorylation by the Aurora B protein kinase [26] (Figure 3). Aurora B phosphorylates several serine/threonine residues in MgcRacGAP *in vitro*, and phosphorylation of Ser387 is essential for its acquisition of RhoGAP activity. Notably, Ser387 is situated in the predicted catalytic domain of MgcRacGAP, suggesting that its phosphorylation directly affects the structure of the catalytic site and thereby influences its ability to interact with particular GTPases.

Phosphorylation of MgcRacGAP is regulated by the cell cycle, and the activation of latent RhoGAP activity by MgcRacGAP is required for cytokinesis. Notably, during cytokinesis, MgcRacGAP localizes to the midbody, which does not require phosphorylation of Ser387, raising the possibility that one (or more) of the other identified phosphorylations of MgcRacGAP mediates its cell-cycle-dependent localization to the midbody. Through such a regulatory mechanism, a single kinase could potentially influence both the substrate specificity and the subcellular localization of a GAP via distinct phosphorylation sites [26].

Phosphorylation-mediated subcellular localization

Several published reports have correlated GAP phosphorylation with a change in protein subcellular localization,

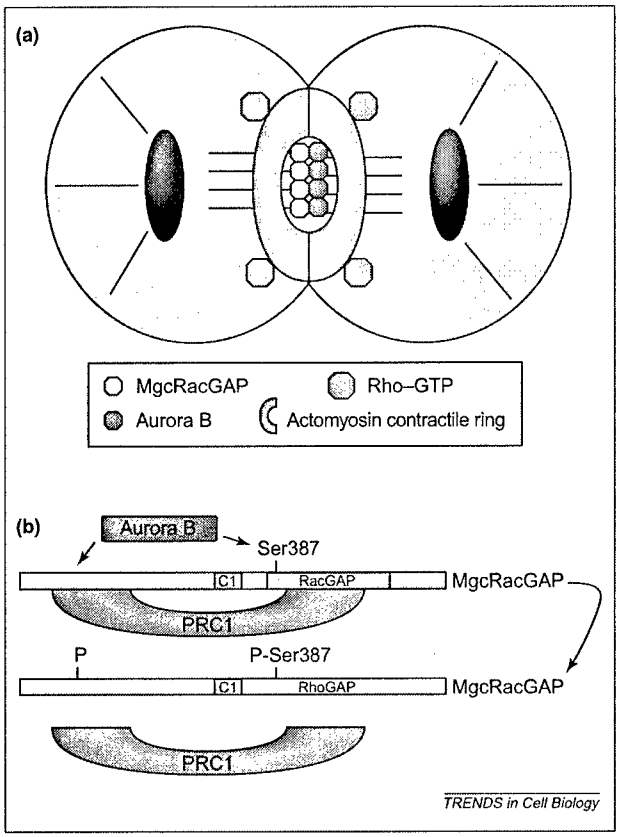


Figure 3. Regulation of MgcRacGAP during cytokinesis. **(a)** Rho–GTP is required for assembly of the contractile actomyosin ring during cytokinesis. MgcRacGAP and the Aurora B kinase colocalize at the spindle midbody. **(b)** Phosphorylation (P) of Ser387 of MgcRacGAP by Aurora B activates the latent RhoGAP activity of MgcRacGAP, resulting in inactivation of Rho and dissociation of the actomyosin ring. Interaction between the spindle-associated protein regulating cytokinesis 1 (PRC1) and MgcRacGAP is also regulated by Aurora B phosphorylation. PRC1 interacts with both the RhoGAP domain and an N-terminal region of MgcRacGAP, and this interaction inhibits the GAP activity of MgcRacGAP.

although the causal role of the phosphorylation has not been firmly established in most cases. For example, casein kinase Iδ phosphorylates ArfGAP1, consequently promoting its interaction with membranes [27]; however, the regulatory mechanism by which this phosphorylation affects protein localization is unknown.

For p190-B RhoGAP, a specific phosphorylation seems to be directly involved in localizing the protein to a membrane subdomain [28]. p190-B RhoGAP has been shown to mediate the inactivation of Rho after exposure of cells to insulin-like growth factor 1 (IGF-1). *In vitro*, insulin receptors can directly phosphorylate p190-B RhoGAP on Tyr306, which is situated in a peptide sequence that closely resembles the consensus phosphorylation site in several known substrates of insulin receptors. It has been shown that exposure of cells to IGF-1 leads to tyrosine phosphorylation of p190-B RhoGAP and induces the rapid translocation of this protein to the lipid-raft-enriched region of the plasma membrane (Figure 4).

Lipid rafts have been identified as membrane subdomains where active Rho–GTP accumulates, suggesting that the redistribution of a RhoGAP to this same subcellular location could be a mechanism to facilitate Rho inactivation. Tyr306 has been established as the requisite phosphorylation site for the IGF-1-induced translocation of p190-B RhoGAP *in vivo*. This tyrosine is located in one of four 'FF domains' of p190-B RhoGAP. FF domains are poorly understood protein interaction motifs present in a relatively few proteins, and the phosphorylation of a tyrosine in this domain might potentially influence an FF-domain-mediated protein interaction with p190-B RhoGAP, thereby directing its accumulation in lipid rafts [28].

Dephosphorylation-mediated GAP regulation

It is worth noting that protein dephosphorylation, via the action of phosphatases, is also likely to have a regulatory role in GAP activity. Although this aspect of GAP regulation has not been widely addressed, evidence suggests that dephosphorylation of p190-B RhoGAP by the SHP-2 tyrosine phosphatase downmodulates its RhoGAP activity [28]. As we describe in the following sections, GAP phosphorylation can influence activity through additional indirect mechanisms, for example, by regulating protein stability and degradation, as well as sensitivity to phospholipid interactions.

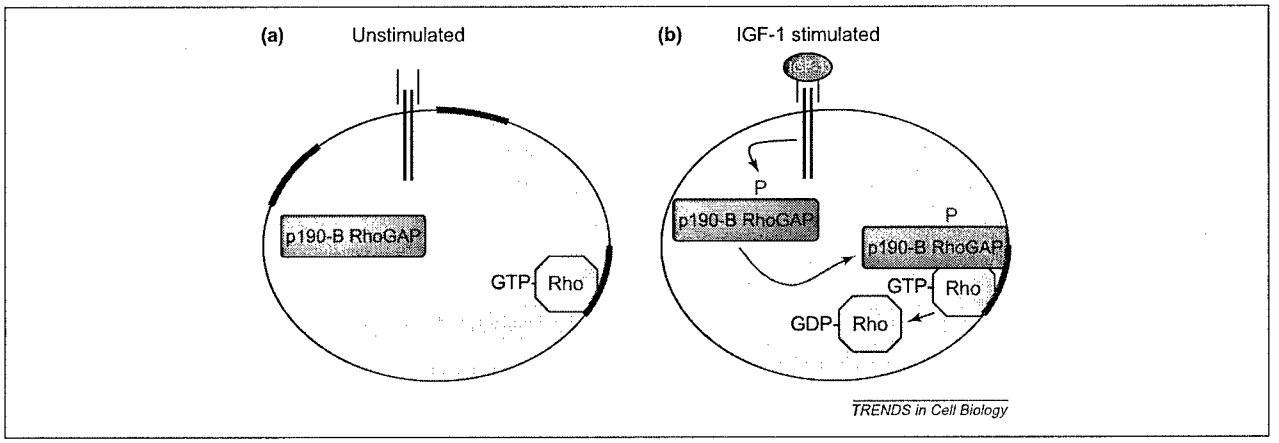


Figure 4. Insulin-like growth factor 1 (IGF-1) promotes the subcellular redistribution of p190-B RhoGAP. In contrast to unstimulated cells **(a)**, when cells are treated with IGF-1 **(b)**, p190-B RhoGAP undergoes tyrosine phosphorylation (P) and consequent translocation to the lipid-raft fraction (dark blue) of the plasma membrane, where active Rho is known to accumulate. Therefore, this phosphorylation event seems to enable p190-B RhoGAP to inactivate Rho effectively.

Regulation by lipids

PHOSPHOLIPIDS and fatty acids are cell membrane components that have been implicated in the regulation of many proteins. In addition to their structural role in maintaining membrane dynamics and fluidity, they can function as docking sites for numerous signaling and trafficking proteins. In addition, several classes of lipid are generated through enzyme activity in response to specific extracellular signals and can thereby function as signal amplifiers.

Many of the prominent physiological lipids seem to have a regulatory role in GTPase-mediated signal transduction. For example, activation of protein kinase C (PKC) by phospholipase Cγ (PLCγ)-mediated generation of diacylglycerol has been implicated in signaling by several of the Ras and Rho family GTPases. Lipids can influence GTPase signaling by affecting various protein components that function as upstream regulators or downstream targets of the GTPase. Moreover, because most small GTPases have covalently bound lipids at their C-terminus and are consequently targeted to membranes, they are located in close proximity to regulators and targets that are potentially influenced by membrane-bound lipids.

Thus, lipids can modulate GTPase-mediated signaling at many levels. In the context of GAP regulation, several studies now point to a likely role for various lipids in the regulation of GAP catalytic function through direct interactions with GAP proteins.

Regulation of RasGAPs

Early studies suggested that lipids have a regulatory role in controlling the two RasGAPs neurofibromin and p120 RasGAP. Those studies showed that various acidic phospholipids and fatty acids have strong inhibitory effects on the catalytic activity of these proteins towards Ras-mediated GTP hydrolysis [29,30]. Subsequent analysis showed that lipid micelles interact directly with the catalytic domain of these RasGAPs and potentially inhibit activity by simply sequestering the protein and reducing its accessibility to its GTPase substrate [31]. The physiological relevance of this observed *in vitro* regulation was not established in those initial studies, and it was proposed that such regulation might not be significant *in vivo*. It was acknowledged, however, that a sequestering mechanism could provide a pathway for regulating cellular GAP activity, possibly by affecting local concentrations of active GAP in membrane subdomains.

Analysis of three mammalian RasGAPs that are structural orthologs of *Drosophila* GAP1 suggests that highly related proteins with identical overall domain structures can be regulated in fundamentally different ways. GAP1-related RasGAPs are characterized by the presence of two phospholipid-binding C2 motifs, followed by a RasGAP catalytic segment and a PH-BTK domain. Analysis of mammalian $GAP1^{m}$ and $GAP1^{IP4BP}$ found the latter to be constitutively associated with the plasma membrane and the former to translocate to the plasma membrane after activation by PI3K [32]. This difference has been attributed to different phosphoinositide-binding specificities of the PH domains of these proteins, which share 63% identity [33].

A third mammalian GAP1 ortholog, termed Ca^{2+}-promoted Ras inactivator (CAPRI), undergoes Ca^{2+}-dependent membrane translocation, which activates its RasGAP activity through an unknown mechanism [34]. The PH domain of CAPRI does not bind phosphoinositides, and the C2 domains of $GAP1^{m}$ and $GAP1^{IP4BP}$ lack residues implicated in Ca^{2+}-dependent lipid binding. Thus, it seems that other closely related GAPs are likely to be regulated through distinct mechanisms.

Regulation of RacGAPs

n-Chimaerin is a Rac- and Cdc42-specific GAP that contains a sequence similar to the phospholipid-dependent phorbol-ester-binding regulatory domain of PKC, suggesting that n-chimaerin might be similarly sensitive to phospholipid regulation. Indeed, results from *in vitro* GAP assays have shown that n-chimaerin activity can be inhibited by some lipids, such as phosphatidylserine and phosphatidic acid, but stimulated by others, including lysophosphatidic acid, phosphatidylinositol (4,5)-bisphosphate [PtdIns(4,5)P_2] and arachadonic acid [35]. Thus, n-chimaerin activity might be highly dependent on both membrane localization and the activation of upstream lipid-generating signals.

Regulation of ArfGAPs

Perhaps the most well-established example of lipid-regulated GAP activity comes from reports that describe the phosphoinositide-dependent activity of the Arf family GAPs. Several of the 26 putative human ArfGAPs show PtdIns(4,5)P_2-dependent or phosphatidylinositol (3,4,5)-trisphosphate [PtdIns(3,4,5)P_3]-dependent GAP activity. For ARAPs, a family of three proteins that contain distinct ArfGAP and RhoGAP catalytic domains, sensitivity to PtdIns(3,4,5)P_3 is mediated by one of five PH domains found in these proteins. Notably, the ArfGAP but not the RhoGAP activity of the ARAPs is dependent on PtdIns(3,4,5)P_3. For ARAP1, it has been shown that the N-terminal PH domain mediates PtdIns(3,4,5)P_3 binding, which suggests that the other domains have distinct regulatory roles, possibly in mediating the interaction of this protein with membranes.

The activation of ARAPs by PtdIns(3,4,5)P_3 indicates that these proteins might be effectors of PI3K signaling. This function has been confirmed for ARAP3, in which a PH domain has been shown to mediate translocation of ARAP3 to the plasma membrane in response to PI3K activation. Some ArfGAPs, such as ArfGAP3, GIT1 and GIT2, do not contain obvious PH domains, but still show phosphoinositide-dependent GAP activity, indicating that another region of the protein mediates lipid regulation. It is also interesting that several of the Arf activators (ArfGEFs), such as ARNO, contain PH domains and show phosphoinositide-induced GEF activity [36].

Taken together, these findings suggest that phosphoinositides have a key role in regulating the cycling of Arf GTPases by stimulating both GEF and GAP activities. Notably, the ArfGAP ASAP1 is directly phosphorylated by the tyrosine kinase Pyk2; this phosphorylation seems to alter the phosphoinositide-binding profile of ASAP1 and, consequently, to inhibit its ArfGAP activity [37]. Such a

finding indicates that phosphorylation might be used more generally to influence the lipid-binding ability of the various GAPs, and thus to influence indirectly lipid-regulated catalytic activity or subcellular localization.

Regulation of RhoGAPs

A recent report demonstrated that the p190 RhoGAPs are also sensitive to phospholipid regulation [38]. The two p190 RhoGAPs (p190-A and p190-B) show catalytic GAP activity toward several Rho family GTPases, including RhoA, Rac1 and Cdc42. Several phospholipids, including phosphatidylserine, PtdIns(4,5)P_2 and phosphatidic acid, are effective inhibitors of p190 RhoGAP activity *in vitro*. Notably, phosphatidylserine and PtdIns(4,5)P_2 have been found to stimulate the GAP activity of p190 RhoGAPs towards the Rac1 GTPase.

Thus, phospholipids can 'switch' the GTPase substrate preference for a GAP, thereby providing a distinct mechanism for GTPase regulation by GAPs. Because phospholipids are not randomly distributed in cellular membranes, such a regulatory mechanism might be used to determine the substrate specificity of a GAP within a particular membrane subdomain. The mechanism by which phospholipids influence the catalytic function of the p190 GAPs, as well as other GAPs, remains rather unclear, and establishing the precise role of these lipids in GAP catalysis will probably require detailed structural studies.

Role of membrane curvature

The BAR domain, which is present in various proteins including several ArfGAPs and RhoGAPs (Table 1), has been recently found to show a crescent-like structure and potentially to function as a sensor of membrane curvature [39]. Thus, GAPs containing BAR domains might be specifically recruited to curved membranes, such as those of budding vesicles.

The assembly of COPI-coated vesicles on Golgi membranes involves the recruitment of the heptameric coatomer complex by Arf1-GTP, whereas COPI coat disassembly before vesicle fusion with target membranes requires ArfGAP1-stimulated hydrolysis of Arf1-GTP. Remarkably, it has been reported that ArfGAP1-stimulated GTP hydrolysis by Arf1 and the subsequent disassembly of the COPI coat increases by more than two orders of magnitude as the curvature of the lipid bilayer approaches that of a transport vesicle. A model that could explain this observation is that the interaction between ArfGAP1 and Arf1 occurs more readily when membrane curvature increases [40]. It is important to note, however, that ArfGAP1 does not contain an obvious BAR domain.

Protein degradation

Protein abundance is an obvious mechanism by which GAP activity can be regulated. In fact, many of the genes identified to encode GAPs are reported to be restricted in their expression, suggesting that cell-type-specific transcriptional elements have a regulatory role; for example, the RacGAP n-chimaerin is expressed specifically in neurons [41]. But other than early studies showing that expression of the RapGAP Spa1 is enhanced by mitogenic stimuli [42], relatively few analyses of the regulation of GAP expression have been reported.

In addition to transcriptional regulation, protein degradation is a potential regulatory mechanism for controlling GAP activity, and several reports have recently confirmed that targeted protein destruction has an apparent role in regulating some GAPs. The RasGAP neurofibromin has been observed to undergo rapid proteolytic degradation via the UBIQUITIN–PROTEASOME SYSTEM upon the treatment of cells with various growth factors [43]. The results indicate that this regulation of neurofibromin can account for an observed increase in the magnitude and duration of Ras activation in response to growth factors. The p120 RasGAP also seems to be regulated by proteolysis; unlike neurofibromin, however, p120 RasGAP is degraded by caspase cleavage, and an N-terminal proteolytic fragment of p120 RasGAP seems to be involved in inhibiting apoptosis [44].

Caspase-independent cleavage of p120 RasGAP in response to infection by coxsackievirus, a small RNA virus, has also been described, and this cleavage seems to correlate with an increase in Ras pathway activity [45]. Notably, p120 RasGAP is a substrate of the herpes simplex virus type 2 protein kinase, IP10, and the expression of this kinase is associated with an increase in Ras activity [46]. Taken together, these findings suggest that viral infection can promote Ras activity through specific degradation or phosphorylation-mediated inhibition of a RasGAP.

The p190-A RhoGAP has been recently shown to undergo targeted destruction via the proteasome pathway [47]. In that study, it was observed that levels of p190-A protein fluctuate during the cell cycle, with a sharp drop observed during late mitosis. It was determined that p190-A undergoes ubiquitination in late mitosis, suggesting that regulation of RhoGAP activity via proteolysis of RhoGAP has a role in the completion of mitosis. Several of the identified RapGAPs, including Tuberin, RapGAP1 and the papillomavirus E6 targeted protein 1 (E6TP1) have been reported to undergo regulated proteolytic degradation [48–50]. Notably, E6TP1 is targeted for degradation by the papillomavirus-encoded E6 protein, and the ability to target E6TP1 is strictly correlated with the ability of the virus to immortalize human mammary epithelial cells [50,51].

In some cases, it seems that specific phosphorylation events provide the signal for targeted degradation of GAPs. For example, phosphorylation of Tuberin by Akt inhibits its association with its binding partner, hamartin (encoded by *TSC1*), leading to its instability and degradation [52]. Similarly, phosphorylation of RapGAP1 by glycogen synthase kinase 3β (GSK-3β) seems to target RapGAP1 for ubiquitin-mediated destruction [49]. Thus, phosphorylation-triggered, protease-mediated degradation of GAPs might be a general mechanism for rapidly inactivating GAP function, presumably as a means of promoting the rapid activation of GTPases.

Concluding remarks

The many members of the Ras superfamily of small GTPases have been implicated in nearly all cellular

processes. It seems that about 0.5% of human genes are likely to encode GAPs for these GTPases, and most of the predicted GAPs contain numerous other functional protein domains and potential phosphorylation sites, indicating that the regulatory control of GTPase activity is highly complex.

Here we have described some of the various regulatory schemes that have been reported to influence the activity of the relatively small subset of GAPs that have been closely examined. Such studies have shown that protein–protein interactions, phosphorylation, lipid interactions, subcellular translocation and targeted degradation all have roles in the modulation of GAP activity; presumably, additional mechanisms will be discovered in future studies.

Considering that a similar number of GEFs seem to be encoded by the human genome, and that GEFs are also frequently associated with various functional domains, it seems likely that a detailed understanding of GTPase regulation will remain a challenge to investigators for many years to come.

Acknowledgements

We apologize to colleagues whose work we could not cite owing to space constraints. Work in our laboratories is supported by grants from the National Institutes of Health and the US Army Medical Research and Materiel Command.

References

1 Takai, Y. *et al.* (2001) Small GTP-binding proteins. *Physiol. Rev.* 81, 153–208
2 Symons, M. and Settleman, J. (2000) Rho family GTPases: more than simple switches. *Trends Cell Biol.* 10, 415–419
3 McCormick, F. (1989) ras GTPase activating protein: signal transmitter and signal terminator. *Cell* 56, 5–8
4 Vincent, S. and Settleman, J. (1999) Inhibition of RhoGAP activity is sufficient for the induction of Rho-mediated actin reorganization. *Eur. J. Cell Biol.* 78, 539–548
5 Bernards, A. (2003) GAPs galore! A survey of putative Ras superfamily GTPase activating proteins in man and Drosophila. *Biochim. Biophys. Acta* 1603, 47–82
6 Hu, F. and Elledge, S.J. (2002) Bub2 is a cell cycle regulated phosphoprotein controlled by multiple checkpoints. *Cell Cycle* 1, 351–355
7 Hu, F. *et al.* (2001) Regulation of the Bub2/Bfa1 GAP complex by Cdc5 and cell cycle checkpoints. *Cell* 107, 655–665
8 Geymonat, M. *et al.* (2003) *In vitro* regulation of budding yeast Bfa1/Bub2 GAP activity by Cdc5. *J. Biol. Chem.* 278, 14591–14594
9 Hofken, T. and Schiebel, E. (2004) Novel regulation of mitotic exit by the Cdc42 effectors Gic1 and Gic2. *J. Cell Biol.* 164, 219–231
10 Li, Y. *et al.* (2004) TSC2: filling the GAP in the mTOR signaling pathway. *Trends Biochem. Sci.* 29, 32–38
11 Nellist, M. *et al.* (2001) TSC2 missense mutations inhibit tuberin phosphorylation and prevent formation of the tuberin–hamartin complex. *Hum. Mol. Genet.* 10, 2889–2898
12 Inoki, K. *et al.* (2002) TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. *Nat. Cell Biol.* 4, 648–657
13 Mochizuki, N. *et al.* (1999) Activation of the ERK/MAPK pathway by an isoform of rap1GAP associated with Gαi. *Nature* 400, 891–894
14 Meng, J. *et al.* (1999) Functional interaction between $G\alpha_z$ and Rap1GAP suggests a novel form of cellular cross-talk. *J. Biol. Chem.* 274, 36663–36669
15 Jordan, J.D. *et al.* (1999) Modulation of rap activity by direct interaction of $G\alpha_o$ with Rap1 GTPase-activating protein. *J. Biol. Chem.* 274, 21507–21510
16 Ban, R. *et al.* (2004) Human mitotic spindle-associated protein PRC1 inhibits MgcRacGAP activity toward Cdc42 during the metaphase. *J. Biol. Chem.* 279, 16394–16402
17 Jenna, S. *et al.* (2002) The activity of the GTPase-activating protein CdGAP is regulated by the endocytic protein intersectin. *J. Biol. Chem.* 277, 6366–6373
18 Wennerberg, K. *et al.* (2003) Rnd proteins function as RhoA antagonists by activating p190 RhoGAP. *Curr. Biol.* 13, 1106–1115
19 Lambert, J.M. *et al.* (2002) Tiam1 mediates Ras activation of Rac by a PI(3)K-independent mechanism. *Nat. Cell Biol.* 4, 621–625
20 Drugan, J.K. *et al.* (2000) The Ras/p120 GTPase-activating protein (GAP) interaction is regulated by the p120 GAP pleckstrin homology domain. *J. Biol. Chem.* 275, 35021–35027
21 Fauchereau, F. *et al.* (2003) The RhoGAP activity of OPHN1, a new F-actin-binding protein, is negatively controlled by its amino-terminal domain. *Mol. Cell. Neurosci.* 23, 574–586
22 Feng, L. *et al.* (2004) PKA phosphorylation and 14–3-3 interaction regulate the function of neurofibromatosis type I tumor suppressor, neurofibromin. *FEBS Lett.* 557, 275–282
23 Okabe, T. *et al.* (2003) RICS, a novel GTPase-activating protein for Cdc42 and Rac1, is involved in the β-catenin-N-cadherin and N-methyl-D-aspartate receptor signaling. *J. Biol. Chem.* 278, 9920–9927
24 Oh, J.S. *et al.* (2004) Regulation of the neuron-specific Ras GTPase-activating protein, synGAP, by Ca^{2+}/calmodulin-dependent protein kinase II. *J. Biol. Chem.* 279, 17980–17988
25 Sano, H. *et al.* (2003) Insulin-stimulated phosphorylation of a Rab GTPase-activating protein regulates GLUT4 translocation. *J. Biol. Chem.* 278, 14599–14602
26 Minoshima, Y. *et al.* (2003) Phosphorylation by aurora B converts MgcRacGAP to a RhoGAP during cytokinesis. *Dev. Cell* 4, 549–560
27 Yu, S. and Roth, M.G. (2002) Casein kinase I regulates membrane binding by ARF GAP1. *Mol. Biol. Cell* 13, 2559–2570
28 Sordella, R. *et al.* (2003) Modulation of Rho GTPase signaling regulates a switch between adipogenesis and myogenesis. *Cell* 113, 147–158
29 Tsai, M.H. *et al.* (1989) The effect of GTPase activating protein upon ras is inhibited by mitogenically responsive lipids. *Science* 243, 522–526
30 Bollag, G. and McCormick, F. (1991) Differential regulation of rasGAP and neurofibromatosis gene product activities. *Nature* 351, 576–579
31 Serth, J. *et al.* (1991) The inhibition of the GTPase activating protein-Ha-ras interaction by acidic lipids is due to physical association of the C-terminal domain of the GTPase activating protein with micellar structures. *EMBO J.* 10, 1325–1330
32 Lockyer, P.J. *et al.* (1997) Distinct subcellular localisations of the putative inositol 1,3,4,5-tetrakisphosphate receptors $GAP1^{IP4BP}$ and $GAP1^{m}$ result from the $GAP1^{IP4BP}$ PH domain directing plasma membrane targeting. *Curr. Biol.* 7, 1007–1010
33 Cozier, G.E. *et al.* (2003) Engineering the phosphoinositide-binding profile of a class I pleckstrin homology domain. *J. Biol. Chem.* 278, 39489–39496
34 Lockyer, P.J. *et al.* (2001) CAPRI regulates Ca^{2+}-dependent inactivation of the Ras-MAPK pathway. *Curr. Biol.* 11, 981–986
35 Ahmed, S. *et al.* (1995) GTPase-activating protein activity of $n_{\alpha 1}$-Chimaerin and effect of lipids. *Methods Enzymol.* 256, 114–125
36 Santy, L.C. *et al.* (1999) Regulation of ARNO nucleotide exchange by a PH domain electrostatic switch. *Curr. Biol.* 9, 1173–1176
37 Kruljac-Letunic, A. *et al.* (2003) The tyrosine kinase Pyk2 regulates Arf1 activity by phosphorylation and inhibition of the Arf-GTPase-activating protein ASAP1. *J. Biol. Chem.* 278, 29560–29570
38 Ligeti, E. *et al.* (2004) Phospholipids can switch the GTPase substrate preference of a GTPase-activating protein. *J. Biol. Chem.* 279, 5055–5058
39 Peter, B.J. *et al.* (2004) BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* 303, 495–499
40 Bigay, J. *et al.* (2003) Lipid packing sensed by ArfGAP1 couples COPI coat disassembly to membrane bilayer curvature. *Nature* 426, 563–566
41 Hall, C. *et al.* (1990) Novel human brain cDNA encoding a 34,000 M_r protein n-chimaerin, related to both the regulatory domain of protein kinase C and BCR, the product of the breakpoint cluster region gene. *J. Mol. Biol.* 211, 11–16
42 Hattori, M. *et al.* (1995) Molecular cloning of a novel mitogen-inducible nuclear protein with a Ran GTPase-activating domain that affects cell cycle progression. *Mol. Cell. Biol.* 15, 552–560
43 Cichowski, K. *et al.* (2003) Dynamic regulation of the Ras pathway via proteolysis of the NF1 tumor suppressor. *Genes Dev.* 17, 449–454

44 Yang, J.Y. and Widmann, C. (2001) Antiapoptotic signaling generated by caspase-induced cleavage of RasGAP. *Mol. Cell. Biol.* 21, 5346–5358
45 Huber, M. *et al.* (1999) Cleavage of RasGAP and phosphorylation of mitogen-activated protein kinase in the course of coxsackievirus B3 replication. *J. Virol.* 73, 3587–3594
46 Smith, C.C. *et al.* (2000) Ras-GAP binding and phosphorylation by herpes simplex virus type 2 RR1 PK (ICP10) and activation of the Ras/MEK/MAPK mitogenic pathway are required for timely onset of virus growth. *J. Virol.* 74, 10417–10429
47 Su, L. *et al.* (2003) $p190^{RhoGAP}$ is cell cycle regulated and affects cytokinesis. *J. Cell Biol.* 163, 571–582
48 Plas, D.R. and Thompson, C.B. (2003) Akt activation promotes degradation of tuberin and FOXO3a via the proteasome. *J. Biol. Chem.* 278, 12361–12366
49 Tsygankova, O.M. *et al.* (2004) Thyroid-stimulating hormone/cAMP and glycogen synthase kinase 3β elicit opposing effects on Rap1GAP stability. *J. Biol. Chem.* 279, 5501–5507
50 Gao, Q. *et al.* (1999) The E6 oncoproteins of high-risk papillomaviruses bind to a novel putative GAP protein, E6TP1, and target it for degradation. *Mol. Cell. Biol.* 19, 733–744
51 Gao, Q. *et al.* (2001) Human papillomavirus type 16 E6-induced degradation of E6TP1 correlates with its ability to immortalize human mammary epithelial cells. *J. Virol.* 75, 4459–4466
52 Dan, H.C. *et al.* (2002) Phosphatidylinositol 3-kinase/Akt pathway regulates tuberous sclerosis tumor suppressor complex by phosphorylation of tuberin. *J. Biol. Chem.* 277, 35364–35370
53 Cichowski, K. and Jacks, T. (2001) NF1 tumor suppressor gene function: narrowing the GAP. *Cell* 104, 593–604
54 Eerola, I. *et al.* (2003) Capillary malformation–arteriovenous malformation, a new clinical and genetic disorder caused by *RASA1* mutations. *Am. J. Hum. Genet.* 73, 1240–1249
55 Henkemeyer, M. *et al.* (1995) Vascular system defects and neuronal apoptosis in mice lacking ras GTPase-activating protein. *Nature* 377, 695–701
56 Faucherre, A. *et al.* (2003) Lowe syndrome protein OCRL1 interacts with Rac GTPase in the trans-Golgi network. *Hum. Mol. Genet.* 12, 2449–2456
57 Billuart, P. *et al.* (1998) Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation. *Nature* 392, 923–926
58 Kutsche, K. *et al.* (2000) Mutations in *ARHGEF6*, encoding a guanine nucleotide exchange factor for Rho GTPases, in patients with X-linked mental retardation. *Nat. Genet.* 26, 247–250
59 Allen, K.M. *et al.* (1998) PAK3 mutation in nonsyndromic X-linked mental retardation. *Nat. Genet.* 20, 25–30
60 Borkhardt, A. *et al.* (2000) The human *GRAF* gene is fused to *MLL* in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9168–9173
61 Wong, C.M. *et al.* (2003) Genetic and epigenetic alterations of *DLC-1* gene in hepatocellular carcinoma. *Cancer Res.* 63, 7646–7651
62 Schaefer, L. *et al.* (1997) Cloning and characterization of a novel rho-type GTPase-activating protein gene (*ARHGAP6*) from the critical region for microphthalmia with linear skin defects. *Genomics* 46, 268–277