Sample Grant Application

Introduction

On the following pages you will find one of the Sample R01 Applications and Summary Statements indexed here:
http://www.niaid.nih.gov/ncn/grants/app/default.htm

Visit the Web site for the most recent information. We may add more in the future.

We are truly indebted to the grantees who've allowed us to post their outstanding applications to help the next generation of investigators write applications.

Copyright

Contact deaweb@niaid.nih.gov with any questions.

PI: **Wahlby, Carolina Ewa Asa**	Title: Image analysis for high-throughput C. elegans infection and metabolism assays	
Received: 02/03/2010	FOA: PA10-067	Council: 10/2010
Competition ID: ADOBE-FORMS-B	FOA Title: Research Project Grant (Parent R01)	
1 R01 GM095672-01	Dual:	Accession Number: 3264072
IPF: 10021177	Organization: BROAD INSTITUTE, INC.	
Former Number:	Department:	
IRG/SRG: MI	AIDS: N	Expedited: N
Subtotal Direct Costs (excludes consortium F&A) Year 1: 250,000 Year 2: 250,000 Year 3: 250,000 Year 4: 250,000 Year 5: 250,000	Animals: N Humans: N Clinical Trial: N Current HS Code: 10 HESC: N	New Investigator: Y Early Stage Investigator: Y
Senior/Key Personnel:	*Organization:*	*Role Category:*
Carolina Wahlby Ph.D.	Broad Institute, Inc.	PD/PI
Anne Carpenter Ph.D	Broad Institute, Inc.	Other Professional-Platform Director
Frederick Ausubel Ph.D	Massachusetts General Hospital	Other (Specify)-Other Significant Contributor
Polina Golland	Massachusetts Institute of Technology	Other (Specify)-Other Significant Contributor
Gary Ruvkun Ph.D	Massachusetts General Hospital	Other (Specify)-Other Significant Contributor
Tamar Riklin-Raviv Ph.D	Massachusetts Institute of Technology	Other (Specify)-Other Significant Contributor

Appendices

2010 wahlby isbi

OMB Number: 4040-0001
Expiration Date: 06/30/2011

APPLICATION FOR FEDERAL ASSISTANCE
SF 424 (R&R)

3. DATE RECEIVED BY STATE	State Application Identifier

1. * TYPE OF SUBMISSION
☐ Pre-application ☒ Application ☐ Changed/Corrected Application

2. DATE SUBMITTED | **Applicant Identifier**

4. a. Federal Identifier

b. Agency Routing Identifier

5. APPLICANT INFORMATION

*** Organizational DUNS:** 623544785

* Legal Name: Broad Institute, Inc.

Department: | Division:

* Street1: 7 Cambridge Center

Street2:

* City: Cambridge | County / Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * ZIP / Postal Code: 02142-1401

Person to be contacted on matters involving this application

Prefix: | * First Name: Stacey | Middle Name:

* Last Name: Donnelly | Suffix:

* Phone Number: 617-714-7130 | Fax Number: 617-714-8972

Email: awards@broadinstitute.org

6. * EMPLOYER IDENTIFICATION *(EIN) or (TIN):* xxxxxxx

7. * TYPE OF APPLICANT: M: Nonprofit with 501C3 IRS Status (Other than Institution of Higher Education)

Other (Specify):

Small Business Organization Type ☐ Women Owned ☐ Socially and Economically Disadvantaged

8. * TYPE OF APPLICATION:
☒ New ☐ Resubmission
☐ Renewal ☐ Continuation ☐ Revision

If Revision, mark appropriate box(es).
☐ A. Increase Award ☐ B. Decrease Award ☐ C. Increase Duration ☐ D. Decrease Duration
☐ E. Other (specify):

* Is this application being submitted to other agencies? Yes ☐ No ☒ What other Agencies?

9. * NAME OF FEDERAL AGENCY: National Institutes of Health

10. CATALOG OF FEDERAL DOMESTIC ASSISTANCE NUMBER:
TITLE:

11. * DESCRIPTIVE TITLE OF APPLICANT'S PROJECT:
Image analysis for high-throughput C. elegans infection and metabolism assays

12. PROPOSED PROJECT:
* Start Date: 12/01/2010 | * Ending Date: 11/30/2015

*** 13. CONGRESSIONAL DISTRICT OF APPLICANT:** MA008

14. PROJECT DIRECTOR/PRINCIPAL INVESTIGATOR CONTACT INFORMATION

Prefix: | * First Name: Carolina | Middle Name:

* Last Name: Wahlby | Suffix: Ph.D.

Position/Title: Computational Biologist

* Organization Name: Broad Institute, Inc.

Department: | Division:

* Street1: 7 Cambridge Center

Street2:

* City: Cambridge | County / Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * ZIP / Postal Code: 02142-1401

* Phone Number: 617-714-7781 | Fax Number:

* Email: carolina@broadinstitute.org

15. ESTIMATED PROJECT FUNDING

a. Total Federal Funds Requested	1,892,525.00
b. Total Non-Federal Funds	0.00
c. Total Federal & Non-Federal Funds	1,892,525.00
d. Estimated Program Income	0.00

16. * IS APPLICATION SUBJECT TO REVIEW BY STATE EXECUTIVE ORDER 12372 PROCESS?

a. YES ☐ THIS PREAPPLICATION/APPLICATION WAS MADE AVAILABLE TO THE STATE EXECUTIVE ORDER 12372 PROCESS FOR REVIEW ON:

DATE:

b. NO ☒ PROGRAM IS NOT COVERED BY E.O. 12372; OR

☐ PROGRAM HAS NOT BEEN SELECTED BY STATE FOR REVIEW

17. By signing this application, I certify (1) to the statements contained in the list of certifications* and (2) that the statements herein are true, complete and accurate to the best of my knowledge. I also provide the required assurances * and agree to comply with any resulting terms if I accept an award. I am aware that any false, fictitious. or fraudulent statements or claims may subject me to criminal, civil, or administrative penalities. (U.S. Code, Title 18, Section 1001)

☒ *** I agree**

**** The list of certifications and assurances, or an Internet site where you may obtain this list, is contained in the announcement or agency specific instructions.***

18. SFLLL or other Explanatory Documentation

Add Attachment | Delete Attachment | View Attachment

19. Authorized Representative

Prefix: * First Name: Stacey Middle Name:

* Last Name: Donnelly Suffix:

* Position/Title: Director, Sponsored Research and Planning

* Organization: Broad Institute, Inc.

Department: Division:

* Street1: 7 Cambridge Center

Street2:

* City: Cambridge County / Parish:

* State: MA: Massachusetts Province:

* Country: USA: UNITED STATES * ZIP / Postal Code: 02142-1401

* Phone Number: 617-714-7130 Fax Number: 617-714-8972

* Email: awards@broadinstitute.org

*** Signature of Authorized Representative**

Scott Breiding

*** Date Signed**

02/03/2010

20. Pre-application

Add Attachment | Delete Attachment | View Attachment

424 R&R and PHS-398 Specific Table Of Contents

Appendix

Number of Attachments in Appendix: 1

OMB Number: 4040-0010
Expiration Date: 08/31/2011

Project/Performance Site Location(s)

Project/Performance Site Primary Location ☐ I am submitting an application as an individual, and not on behalf of a company, state, local or tribal government, academia, or other type of organization.

Organization Name: Broad Institute, Inc.

DUNS Number: 6235447850000

* Street1: 7 Cambridge Center

Street2:

* City: Cambridge County:

* State: MA: Massachusetts

Province:

* Country: USA: UNITED STATES

* ZIP / Postal Code: 02142-1401 * Project/ Performance Site Congressional District: MA 008

Project/Performance Site Location 1 ☐ I am submitting an application as an individual, and not on behalf of a company, state, local or tribal government, academia, or other type of organization.

Organization Name:

DUNS Number:

* Street1:

Street2:

* City: County:

* State:

Province:

* Country: USA: UNITED STATES

* ZIP / Postal Code: * Project/ Performance Site Congressional District:

Additional Location(s) Add Attachment Delete Attachment View Attachment

RESEARCH & RELATED Other Project Information

1. * Are Human Subjects Involved? ☐ Yes ☒ No

1.a If YES to Human Subjects

Is the Project Exempt from Federal regulations? ☐ Yes ☐ No

If yes, check appropriate exemption number. ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6

If no, is the IRB review Pending? ☐ Yes ☐ No

IRB Approval Date:

Human Subject Assurance Number:

2. * Are Vertebrate Animals Used? ☐ Yes ☒ No

2.a. If YES to Vertebrate Animals

Is the IACUC review Pending? ☐ Yes ☐ No

IACUC Approval Date:

Animal Welfare Assurance Number

3. * Is proprietary/privileged information included in the application? ☐ Yes ☒ No

4.a. * Does this project have an actual or potential impact on the environment? ☐ Yes ☒ No

4.b. If yes, please explain:

4.c. If this project has an actual or potential impact on the environment, has an exemption been authorized or an environmental assessment (EA) or environmental impact statement (EIS) been performed? ☐ Yes ☐ No

4.d. If yes, please explain:

5. * Is the research performance site designated, or eligible to be designated, as a historic place? ☐ Yes ☒ No

5.a. If yes, please explain:

6. * Does this project involve activities outside of the United States or partnerships with international collaborators? ☐ Yes ☒ No

6.a. If yes, identify countries:

6.b. Optional Explanation:

7. * Project Summary/Abstract 1240-Abstract.pdf Add Attachment Delete Attachment View Attachment

8. * Project Narrative 1241-Narrative.pdf Add Attachment Delete Attachment View Attachment

9. Bibliography & References Cited 1242-BibliographyResubmitted.pdf Add Attachment Delete Attachment View Attachment

10. Facilities & Other Resources 1243-FacilitiesAndResources.pdf Add Attachment Delete Attachment View Attachment

11. Equipment 1244-Equipment.pdf Add Attachment Delete Attachment View Attachment

12. Other Attachments Add Attachments Delete Attachments View Attachments ☐

Abstract

High-throughput screening (HTS) is a technique for searching large libraries of chemical or genetic perturbants, to find new treatments for a disease or to better understand disease pathways. As automated image analysis for cultured cells has improved, microscopy has emerged as one of the most powerful and informative ways to analyze screening samples. However, many diseases and biological pathways can be better studied in whole animals—particularly diseases that involve organ systems and multicellular interactions, such as metabolism and infection. The worm *Caenorhabditis elegans* is a well-established and effective model organism, used by thousands of researchers worldwide to study complex biological processes. Samples of *C. elegans* can be robotically prepared and imaged by high-throughput microscopy, but existing image-analysis methods are insufficient for most assays. In this project, image-analysis algorithms that are capable of scoring high-throughput assays of *C. elegans* will be developed.

The algorithms will be tested and refined in three high-throughput screens, which will uncover chemical and genetic regulators of fat metabolism and infection: (1) A *C. elegans* viability assay to identify modulators of infection. The proposed algorithms use a probabilistic shape model of *C. elegans* in order to identify and measure individual worms even when the animals touch or cross. These methods are the basis for quantifying many other phenotypes, including body morphology and subtle variations in reporter signal levels. (2) A *C. elegans* lipid assay to identify genes that regulate fat metabolism. The algorithms proposed for illumination correction, level-set–based foreground segmentation, well-edge detection, and artifact removal will result in improved robustness in high-throughput experiments. (3) A fluorescence gene expression assay to identify regulators of the response of the *C. elegans* host to *Staphylococcus aureus* infection. The proposed techniques for constructing anatomical maps of *C. elegans* will make it possible to quantify a variety of changes in fluorescent localization patterns in a biologically relevant way.

In addition to discovering new metabolism- and infection-related drugs and genetic regulators through these specific screens, this work will provide the *C. elegans* community with (a) a new framework for extracting morphological features from *C. elegans* for quantitative analysis of this organism, and (b) a versatile, modular, open-source toolbox of algorithms enabling the discovery of genetic pathways, chemical probes, and drug candidates in whole organism high-throughput screens relevant to a variety of diseases.

This work is a close collaboration with *C. elegans* experts Fred Ausubel and Gary Ruvkun at Massachusetts General Hospital/Harvard Medical School, with Polina Golland and Tammy Riklin-Raviv, experts in model-based segmentation and statistical image analysis at MIT's Computer Science and Artificial Intelligence Laboratory, and with Anne Carpenter, developer of open-source image analysis software at the Broad Institute.

Public Health Relevance/Narrative

Large-scale screening experiments that test the effects of thousands of chemicals or genetic perturbants by microscopy and image analysis can discover new treatments and help biomedical scientists understand disease mechanisms. Microscopy screens of cultured cells are routine, but researchers wish to study complex processes like metabolism and infection in a whole animal like the tiny worm *Caenorhabditis elegans*, for which existing image analysis methods are insufficient. The goal of this research is to develop open-source software to automatically identify and measure *C. elegans* in microscopy images, thereby making it possible for researchers worldwide to screen a wide variety of complex biological processes related to human disease.

Facilities and Resources

The work proposed in this application is entirely computational. Although the project team will visit the other collaborators' laboratories for meetings and to provide guidance on sample preparation and image acquisition, the vast majority of the work will take place in offices shared with the Imaging Platform on the sixth floor of the Broad Institute's building at 7 Cambridge Center in the Kendall Square/MIT area of Cambridge, Massachusetts. This building is across the street from the Computer Sciences/Artificial Intelligence Laboratoy at the Massachusetts Institute of Technology (where collaborators Golland and Riklin-Raviv work) and a fifteen-minute walk from Massachusetts General Hospital (MGH), the location of the *C. elegans* laboratories. Our collaborators, the Ausubel and Ruvkun groups, will be performing their separately-funded sample preparation, imaging, and biological followup work at MGH, in their laboratories and in the *C. elegans* High-Throughput Screening Core Facility.

Imaging Platform of the Broad Institute of Harvard and MIT

The budget includes funding for 240 hours/year of Imaging Platform time. The Imaging Platform is set up as part of the Broad Institute's Specialized Service Facility and offers services related to image-analysis software. Dr. Anne Carpenter, key personnel for this proposal, directs this established and well-integrated team, and its role in the project will be to take the novel algorithms developed by the main project team and implement and disseminate these algorithms to the biological community via polished end-user software. Specifically, the team will add the algorithms to the open-source image-analysis software project CellProfiler, which was launched in 2005 and has rapidly become the world standard for open-source high-throughput image analysis.

Hiring this team to provide these services allows several specialists to be dedicated to the tasks within their expertise without requiring hiring and training several people specifically for this project. We find that our team of extraordinarily productive professionals is lean and efficient as compared to equivalent funding spent for PhD students [redacted] who in any case cannot be hired by the Broad except by someone with a co-existing faculty appointment at Harvard and MIT, since we are not a degree-granting institution.

The team's expertise covers image assay development, algorithm development, software engineering, and software tutorial administration:

Thouis R. Jones, Ph.D. (Computational Biologist). After almost a decade of software development in industry, Dr. Jones went on to earn a Ph.D. in computer science from MIT. As part of this work, he co-founded the CellProfiler project with Dr. Carpenter and developed its underlying algorithms. He has spent the past six years developing advanced image analysis and data mining algorithms for image-based screens, including the iterative machine-learning methods described in this proposal. Based on his experience with cell-based imaging screens, he will contribute to algorithm development for *C. elegans* image analysis and data mining, especially machine learning, illumination correction, and thresholding. He will also provide scientific and software engineering guidance to the Imaging Platform team.

Vebjorn Ljosa, Ph.D. (Computational Biologist). Dr. Ljosa earned his Ph.D. in computer science from the University of California, Santa Barbara. He is a senior member of the Imaging Platform, and has five years of experience in developing analysis methods for biological image sets, including segmentation of in-vivo neurons, phenotype discovery, dimensionality reduction, large-scale machine learning for recognizing subtle phenotypes, and most recently, algorithms for *C. elegans*. Dr. Ljosa will be responsible for machine-learning methods to distinguish live and dead worms (Aim 1), select distinguishing features (Aim 2), and classify worms by localization pattern (Aim 3).

Lee Kamentsky (Software Engineer). With nearly 30 years of professional software engineering experience, including writing the image-analysis and hardware-control software for CompuCyte's Laser Scanning Cytometer, Lee Kamentsky joined the team in 2008. He is the lead software developer for the CellProfiler project. For this project he will implement and optimize *C. elegans* algorithms, ensure documentation and unit testing of the code, and guide other team members who contribute to the code, including code review.

Mark Bray, Ph.D. (Computational Biologist). Dr. Bray joined the team in 2008 after completing a postdoctoral fellowship at Harvard. Dr. Bray's primary responsibility is image assay development: he will assist Dr. Wahlby and the *C. elegans* collaborators on biological validation for the *C. elegans* software. He will also manage the public dissemination of images produced for this proposal as well as the software tools. Dr. Bray leads existing efforts for CellProfiler outreach; for this project, he will develop and deliver *C. elegans*-specific curriculum for in-person and online tutorials and training sessions and answer questions from the public online forum. During

years 3 through 5, he will present tutorials or workshops at the Broad Institute and at *C. elegans* meetings as part of our proposal to disseminate the software to the scientific community.

Margaret Anthony (IT Administrator) Ms. Anthony handles transferring and accessing images from collaborators for this project, managing the computational infrastructure for the software distributed by the team, as well as their electronic lab notebook. She will also set up and manage online registration and other arrangements for training sessions for the *C. elegans* software developed by the team.

Environment at the Broad Institute of MIT and Harvard

Focus on biomedical technology: The structure and mission of the Broad Institute is unusual. Its buildings contain six faculty laboratories and eight technology platforms. The technology platforms are teams of professional scientists who focus on the discovery, development, and optimization of the critical technological tools needed to obtain and analyze massive amounts of genome-related data. Platform scientists have the expertise and organization to carry out major projects that could not be done within a single research laboratory, and work closely with the scientific programs and collaborators around the world to tackle critical questions in human biology and disease. The Broad Platforms are: the Imaging Platform (described above), Biological Samples, Chemical Biology, Genome Sequencing, Genetic Analysis, Metabolite Profiling, Proteomics, and RNAi.

At the same time, the Broad's scientific programs nucleate more than 150 biomedical faculty who are Associate Members of the Broad, including Drs. Ausubel, Golland, and Ruvkun. While their laboratories are located outside the building, Associate Members and their graduate students and postdocs meet at the Broad out of shared commitment to critical biomedical research areas: Cancer, Cell Circuits, Chemical Biology, Computational Biology and Bioinformatics, Epigenomics, Genome Biology, Infectious Disease, Medical and Population Genetics, Metabolism, and Psychiatric Disease. This provides a rich collaborative environment dedicated to clinical application of biomedical research.

Focus on high-throughput experimentation: Particularly relevant are the RNAi and Chemical Biology Platforms: both are focused on high-throughput screening, and the latter is a site of the Molecular Libraries Probe Production Centers Network (MLPCN). This provides an intellectual environment immersed in designing and interpreting high-throughput experiments, which has been very helpful so far in developing the *C. elegans* projects described in the proposal.

Focus on the productive interaction between computation and biomedicine: Important to this proposal is the unique collaborative environment and the outstanding scientific expertise at the Broad Institute: more than 1,000 scientists with backgrounds in clinical medicine, cancer biology, molecular biology, statistical genetics, engineering, applied mathematics, physics and computational biology. In addition, there are more than 75 full-time software engineers helping to develop systems to acquire, manage, analyze and share data. This team-based approach has ensured that computational efforts are driven by biological questions, and that biologist users can quickly and critically evaluate computational solutions. This interchange is facilitated by several seminars each day on various biomedical and computational topics, including infectious disease, metabolism, software engineering and computational biology.

Productive collaborations: For this particular project, the proximity of the Wahlby group to the collaborating laboratories is key to productive collaboration. In addition to Dr. Wahlby sharing offices with the Imaging Platform, the collaborating groups meet face-to-face often; Drs. Ausubel, Golland, and Ruvkun are all Associate Members of the Broad and members of their laboratories are often in the building attending seminars, using equipment, or interacting with members of the project team. The proximity of the Broad to the rich intellectual resources of MIT, Harvard, and other local institutions also offers unparalleled opportunities to draw upon the intellectual resources of researchers in the area, as needed. For example, colleagues of Dr. Wahlby have begun exploring large-scale machine-learning methods, as recently developed nearby at MIT, as an approach to overcoming the otherwise intractably large data sets produced by cellular imaging experiments.

Early-stage investigator support

Dr. Wahlby is exempt from teaching responsibilities at the Broad Institute and her salary is paid primarily from startup funds as she establishes her research group. She is provided the office space required for herself and members of her group. Dr. Wahlby also receives funds for traveling to at least one scientific conference per year, and she retains strong ties to the Centre for Digital Image Analysis at Uppsala University where she maintains

an appointment as Associate Professor and oversees 2 graduate students.

- **Training:** The Broad Institute invests in the success of its early-stage investigators by providing various series of short-term classes, workshops, and seminars. In addition to the scientific seminars in biomedicine, computation, and software engineering described above, topics include management, grant-writing, and career development. Dr. Wahlby has already taken advantage of many of these opportunities as she establishes her research agenda. She is also provided with tuition reimbursement for external courses in scientific areas as well as management, both of which are readily available at dozens of institutions in the area, including MIT across the street.

- **Mentoring:** Mentoring is also heavily emphasized at the Broad, and suitable mentors abound. Dr. Wahlby is adjacent to Broad faculty member Dr. Aviv Regev, recent winner of the Burroughs Wellcome Fund Career Award at the Scientific Interface, an NIH Pioneer Award, and an HHMI Early Career Award. Dr. Jill Mesirov, also at the Broad Institute, is another resource, particularly for algorithm and software development. Dr. Wahlby has so far received her most direct mentoring from Dr. Anne Carpenter, Director of the Imaging Platform, who introduced her to the challenges and rewards of the *C. elegans* projects and provides ongoing guidance in establishing a research group.

- **Facilities and administration support:** The Broad Institute provides full supporting services that enable researchers to focus almost entirely on research. Facilities, cleaning, and maintenance are professionally managed. Logistical support is provided by an administration committed to ensuring that the Broad's organizational infrastructure is completely transparent and supportive of the Broad's scientific mission. This includes assisting investigators in following best practices and complying with all relevant regulations and policies. The Finance team supports financial needs through services of professionals skilled in business development, controller functions, sponsored research, cost analysis, procurement, administrative management, and budgeting.

- In short, the Broad Institute provides a non-heirarchical environment in which early-stage investigators can thrive on the strength of their ideas and their ability to convert those ideas into biomedical discoveries.

C. elegans resources

Ausubel and Ruvkun laboratories: The *C. elegans* groups have the equipment and necessary resources to carry out the work that will interface with this proposal. Both laboratories are located in the newly constructed (2005) Richard B. Simches Research Center on the Massachusetts General Hospital main campus. Together, the laboratories have benches and attached desks for roughly 30 researchers as well as office space for bioinformaticians. The Ausubel laboratory has two shared rooms of BL-2 approved research space. The Simches building has a full complement of shared general and specific state-of-the-art laboratory facilities available for use by the Ausubel and Ruvkun laboratories including instrument rooms, a Drosophila growth room, cold rooms, microscope rooms, a walk-in plant growth facility, a balance and chemical storage room, a media preparation facility, a glass washing/sterilization facility, a tissue culture facility, electrophoresis and gel rooms, a dark booth and a specialty dark room as well as laboratory supply storage space. From a broader perspective, the Simches Research Building is designed around thematic centers that house many laboratories at MGH at the forefront of modern biological research. Thus in addition to the Department of Molecular Biology, the Simches building has multi-investigator groups studying systems biology, human genetics, stem cells, chemical genetics, genomics, and computational and integrative biology. The 8th floor houses a core facility for electron microscopy and immunoelectron microscopy. Lab members each have their own computer.

***C. elegans* High Throughput Screening Core Facility:** A resource supporting the *C. elegans* collaborators on the proposed project is their *C. elegans* high throughput screening core facility, operated jointly by the Ausubel and Ruvkun laboratories on the 7th floor of the Simches building. The *C. elegans* Core Facility provides guidance to take a manual *C. elegans* assay and develop it into an automated, high throughput, high content screen, and then, makes available the equipment and technical help to carry out the screen. The staff of the *C. elegans* Core Facility includes a Ph.D. level screen manager, a research technician, and a bioinformatics specialist. The following key components are available in the *C. elegans* Core Facility:

- Titertek MapC2 Liquid and Agar Dispenser: Custom made liquid dispenser that was adapted with heating modules, insulated tubing and slow pump mechanism to enable dispensing of viscous agar.
- Union Biometrica Copas BioSort: The BioSort worm-handling robot automates analysis, sorting and dispensing of "large" objects such as *C. elegans* using object size and intensity of fluorescent markers.
- Zeiss Axioskop 2.0 upgraded to an automated image acquisition system: This microscope has been retrofitted with an XYZ Proscan stage, shutters, and filter wheels for automation. The imaging system includes a high sensitivity Qimaging Retiga EXi SVGA high-speed cooled digital camera, optimized fluorescent filter sets from Chroma, and a Ec Plan Neofluar 2.5x low power objective. The system is controlled by Surveyor acquisition software with a Turboscan kit and an OASIS-blue joystick controller.
- Thermo Scientific Forma Model 3940 Environmental Chamber with humidity control.
- Data storage and analysis infrastructure: Dedicated Dell PowerEdge 2950 server with two dualcore 3.73GHZ CPUs, six 300GB SAS hard drives, and 16GB RAM. Dell MD1000 disk array of fifteen 1TB hard drives which is attached to the PowerEdge 2950 server via an SCSI cable. Structured workflow from data acquisition to analysis is carried out on the dedicated server. The data outflow is automatically routed into a custom Oracle database with a user-friendly web-based data interface.
- The Department of Molecular Biology has dedicated, full-time core Computer Services staff as well as a new Bioinformatics Core Facility consisting of three full-time bioinformaticians.

Equipment

As noted in the Facilities section, the work proposed in this application is entirely computational. The project's budget includes fees on a pro-rated FTE basis for provision and maintenance of a personal computer for each team member. The Broad Institute has a state-of-the-art information technology infrastructure and a team of dedicated staff to professionally manage it. The data storage hardware and the cluster computing nodes purchased for this project will be incorporated into this infrastructure, described as follows:

More than 4,000 sq. ft. of space is dedicated to computer server rooms at the Broad. The principal computing resource is a large compute farm containing both IBM Bladecenter and Dell M1000 hardware comprising more than 650 Linux nodes with more than 2400 processor cores. Shared access to the compute farm is managed by Platform Computing's Load Sharing Facility (LSF). In addition to the compute farm, 200 servers run Linux or Windows dedicated to specific applications, of which twelve are large memory (64–256 GB) servers running Linux for memory-intensive applications such as whole-genome assembly. File storage is provided by network-attached storage products from Network Appliance, Sun Microsystems, and Isilon Systems, which collectively offer access to over 2.8 PB of usable file space via NFS and CIFS protocols. Current SAN storage capacity is over 120 TB, hosted on EMC Clariion and Data Direct Networks products. The primary database environment is Oracle 10g, with approximately 90 TB of SAN-attached storage currently allocated across 36 Oracle servers. Production and research instances of MySQL databases are also supported. Backup and archiving of data is done between the two main Broad buildings so that the backup data is always in a separate facility from the primary data to provide disaster recovery capability. Backup management and performance is handled by EMC's Legato software in conjunction with a 200 TB FalconStor-based virtual tape library, an IBM enterprise-class LTO tape library, and an HP SDLT tape library. All desktop and laptop computers are backed up through a dedicated system served by Atempo software. Cisco and Force10 network switches provide 1 Gbps connectivity throughout the buildings, with a 10 Gbps infrastructure providing server-to-server and server-to-storage connectivity and a 20 Gbps inter-building link. The external link to the Internet is 1 Gbps. All production services are monitored 24x7. The server and storage infrastructure is protected by a UPS and a diesel generator that provides backup power for both the computational hardware and the associated cooling systems. An IT staff of 35 employees provides support for all computing resources, including production and research applications.

OMB Number: 4040-0001
Expiration Date: 06/30/2011

RESEARCH & RELATED Senior/Key Person Profile (Expanded)

PROFILE - Project Director/Principal Investigator

Prefix: | * First Name: Carolina | Middle Name:
* Last Name: Wahlby | Suffix: Ph.D.
Position/Title: Computational Biologist | Department:
Organization Name: Broad Institute, Inc. | Division:
* Street1: 7 Cambridge Center
Street2:
* City: Cambridge | County/ Parish:
* State: MA: Massachusetts | Province:
* Country: USA: UNITED STATES | * Zip / Postal Code: 02142-1401
* Phone Number: 617-714-7781 | Fax Number:
* E-Mail: carolina@broadinstitute.org

Credential, e.g., agency login: xxxxxxx

*** Project Role:** PD/PI | **Other Project Role Category:**

Degree Type: Ph.D.
Degree Year: 2003

***Attach Biographical Sketch** 1234-WahlbyBiosketch.pdf [Add Attachment] [Delete Attachment] [View Attachment]
Attach Current & Pending Support [Add Attachment] [Delete Attachment] [View Attachment]

PROFILE - Senior/Key Person 4

Prefix: | * First Name: Anne | Middle Name: E.
* Last Name: Carpenter | Suffix: Ph.D
Position/Title: Director, Imaging Platform | Department:
Organization Name: Broad Institute, Inc. | Division:
* Street1: 7 Cambridge Center
Street2:
* City: Cambridge | County/ Parish:
* State: MA: Massachusetts | Province:
* Country: USA: UNITED STATES | * Zip / Postal Code: 02142-1401
* Phone Number: 617-714-7750 | Fax Number:
* E-Mail: anne@broadinstitute.org

Credential, e.g., agency login: xxxxxxx

*** Project Role:** Other Professional | **Other Project Role Category:** Platform Director

Degree Type: Ph.D.
Degree Year: 2003

***Attach Biographical Sketch** 1235-CarpenterBiosketch.pdf [Add Attachment] [Delete Attachment] [View Attachment]
Attach Current & Pending Support [Add Attachment] [Delete Attachment] [View Attachment]

RESEARCH & RELATED Senior/Key Person Profile (Expanded)

PROFILE - Senior/Key Person 2

Prefix: | * First Name: Frederick | Middle Name: M.

* Last Name: Ausubel | Suffix: Ph.D

Position/Title: Professor of Genetics | Department:

Organization Name: Massachusetts General Hospital | Division:

* Street1: 185 Cambridge Street

Street2: Simches Research Ctr CPZN# 7808

* City: Boston | County/ Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * Zip / Postal Code: 02114-2790

* Phone Number: 617-726-5969 | Fax Number:

* E-Mail: ausubel@molbio.mgh.harvard.edu

Credential, e.g., agency login: xxxxxxx

*** Project Role:** Other (Specify) | **Other Project Role Category:** Other Significant Contributor

Degree Type: Ph.D.

Degree Year: 1972

***Attach Biographical Sketch** 1236-AusubelNIHBiosketch.pdf | Add Attachment | Delete Attachment | View Attachment

Attach Current & Pending Support | Add Attachment | Delete Attachment | View Attachment

PROFILE - Senior/Key Person 3

Prefix: | * First Name: Polina | Middle Name:

* Last Name: Golland | Suffix:

Position/Title: Associate Professor | Department:

Organization Name: Massachusetts Institute of Technology | Division:

* Street1: 32 Vassar Street

Street2: 32-D470

* City: Cambridge | County/ Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * Zip / Postal Code: 02139-4390

* Phone Number: 617-253-8005 | Fax Number:

* E-Mail: polina@csail.mit.edu

Credential, e.g., agency login: xxxxxxx

*** Project Role:** Other (Specify) | **Other Project Role Category:** Other Significant Contributor

Degree Type: Ph.D.

Degree Year: 2001

***Attach Biographical Sketch** 1237-Golland_Biosketch_Broad | Add Attachment | Delete Attachment | View Attachment

Attach Current & Pending Support | Add Attachment | Delete Attachment | View Attachment

RESEARCH & RELATED Senior/Key Person Profile (Expanded)

PROFILE - Senior/Key Person 4

Prefix: | * First Name: Gary | Middle Name:

* Last Name: Ruvkun | Suffix: Ph.D

Position/Title: Molecular Biologist | Department:

Organization Name: Massachusetts General Hospital | Division:

* Street1: 185 Cambridge Street

Street2: Richard B. Simches Research Center

* City: Boston | County/ Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * Zip / Postal Code: 02114-2790

* Phone Number: 617-726-5959 | Fax Number:

* E-Mail: ruvkun@frodo.mgh.harvard.edu

Credential, e.g., agency login: xxxxxxx

*** Project Role:** Other (Specify) | **Other Project Role Category:** Other Significant Contributor

Degree Type: Ph.D.

Degree Year: 1982

***Attach Biographical Sketch** 1238-Ruvkunbiosketch.pdf [Add Attachment] [Delete Attachment] [View Attachment]

Attach Current & Pending Support [Add Attachment] [Delete Attachment] [View Attachment]

PROFILE - Senior/Key Person 5

Prefix: | * First Name: Tamar | Middle Name:

* Last Name: Riklin-Raviv | Suffix: Ph.D

Position/Title: Postdoctoral Researcher | Department:

Organization Name: Massachusetts Institute of Technology | Division:

* Street1: 32 Vasser Street

Street2: 32-D4320

* City: Cambridge | County/ Parish:

* State: MA: Massachusetts | Province:

* Country: USA: UNITED STATES | * Zip / Postal Code: 02139-4390

* Phone Number: 617-253-2986 | Fax Number:

* E-Mail: tammy@CSAIL.MIT.EDU

Credential, e.g., agency login: xxxxxxx

*** Project Role:** Other (Specify) | **Other Project Role Category:** Other Significant Contributor

Degree Type: Ph.D.

Degree Year: 2008

***Attach Biographical Sketch** 1239-RiklinRavivBiosketch.pdf [Add Attachment] [Delete Attachment] [View Attachment]

Attach Current & Pending Support [Add Attachment] [Delete Attachment] [View Attachment]

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Wählby, Carolina	Computational Biologist Broad Institute of Harvard and MIT
eRA COMMONS USER NAME (credential, e.g., agency login) XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Uppsala University, Uppsala, Sweden (Thesis work at the Karolinska Institute, Stockholm, Sweden)	M.Sc.	06/98	Molecular Biotechnology
Centre for Image Analysis, Uppsala University, Uppsala, Sweden	Ph.D.	10/03	Digital Image Analysis
Dept. Genetics and Pathology, Uppsala University, Uppsala, Sweden	postdoctoral	06/08	Applied image analysis in molecular medicine

A. Personal Statement

The goal of the proposed research is to develop image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism. My research is focused on the development of algorithms for image analysis applied to biological images acquired by microscopy; the questions addressed have ranged from various types of quantification and mapping of fluorescent reporters in cells and tissue to single molecule classification on microarrays. My strongest expertise lies in image segmentation, which is a crucial initial step in every kind of image analysis task, before quantitative information can be robustly extracted from the objects of interest. The algorithms I developed for cell segmentation have served as a foundation for the open-source software CellProfiler and have proven very robust in high-throughput applications. At the Department of Genetics and Pathology, my postdoctoral research was very application-oriented, giving me broad experience in microscopy and issues related to sample preparation and handling in relation to image quality and analysis.

As project manager for the image analysis part of the EU-financed ENLIGHT project I successfully managed researchers from academia as well as industry in 5 different European countries, producing a number of peer-reviewed collaborative publications. This highly interdisciplinary project taught me the importance of communication between collaborators from different fields of expertise, which will be of importance for the proposed project. My experience from supervising PhD students as assistant and associate professor at Uppsala University has also helped me develop skills for project planning, execution and dissemination. My most recent work at the Imaging Platform of the Broad Institute has given me insight about many aspects related to high-throughput experiments, and the expertise available at the platform provides a strong support for the proposed project.

Initial work on *C. elegans* images in collaboration with my co-investigators Polina Golland and Tammy Riklin-Raviv at M.I.T. Computer Science and Artificial Intelligence Lab has already led to a peer-reviewed paper (Wählby et al., 2010) describing methods applied to data provided by our main collaborators Gary Ruvkun and Fred Ausubel at Massachusetts General Hospital/ Harvard Medical School: a well established collaboration initiated by my co-investigator Anne Carpenter, leading to the publication of the first automated image-based screen on *C.elegans* (Moy et al., ACS Chemical Biology, 2009).

To conclude, my strong computational experience in combination with experience from collaborative interdisciplinary projects together with well established contacts provides a basis for successful project leadership.

B. Positions and Honors

Positions:

1998-2003 PhD Student, Uppsala University, Sweden, Centre for Image Analysis, supervised by professor Ewert Bengtsson
2004-2009 Assistant Professor, the Centre for Image Analysis, Uppsala University, Sweden (part-time from 2005)
2005-2008 Postdoctoral Fellow / researcher (part time) at the Dept. Genetics and Pathology, Molecular Medicine, Laboratory of Ulf Landegren, MD, PhD, and Mats Nilsson, PhD.
2009-present Associate Professor, Centre for Image Analysis, Uppsala University, Sweden.
2009-present Computational Biologist, Broad Institute of Harvard and MIT

Other Experience and Professional Memberships

International Association for Pattern Recognition (IAPR), since 1998, Swedish Society for Automated Image Analysis (SSBA), since 1998, International Society for Analytical Cytology (ISAC), since 2001; on the EU Project ENLIGHT Program Management Board, since 2006.

C. Selected peer-reviewed publications (chosen from 24)

Most relevant to the current application

1. **Wählby C**, Riklin-Raviv T, Ljosa V, Conery AL, Golland P, Ausubel FM, Carpenter AE (2010) Resolving clustered worms via probablistic shape models. Accepted for publication in proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI) 2010.
2. Allalou A, Pinidiyaarachchi A, **Wählby C**. Robust signal detection in 3D fluorescence microscopy. Cytometry A. 2009 Apr;94(1):58-65. PMID: 19760746
3. **Wählby C**, Sintorn IM, Erlandsson F, Borgefors G, Bengtsson E. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. J Microsc. 2004 Jul;215(Pt1):67-76. PMID: 15230877
4. **Wählby C**, Lindblad J, Vondrus M, Bengtsson E, Björkesten L. Algorithms for cytoplasm segmentation of fluorescence labelled cells. Anal Cell Pathol. 2002;24(2-3):101-11. PMID: 12446959
5. Gavrilovic M, **Wählby C**. Quantification of colocalization and cross-talk based on spectral angles. J Microsc. 2009 Jun;234(3):311-324. PMID: 19493110

Additional recent publications of importance to the field

6. Erlandsson F, **Linnman (-Wählby) C**, Ekholm S, Bengtsson E, Zetterberg A. A detailed analysis of cyclin A accumulation at the G(1)/S border in normal and transformed cells. Exp Cell Res. 2000 A 25;259(1):86-95. PMID: 10942581
7. **Wählby C**, Erlandsson F, Bengtsson E, Zetterberg A. Sequential immunofluorescence staining and image analysis for detection of large numbers of antigens in individual cell nuclei. Cytometry. 2002 Jan 1;47(1):32-41. PMID: 11774347
8. Erlandsson F, **Wählby C**, Ekholm-Reed S, Hellström AC, Bengtsson E, Zetterberg A. Abnormal expression pattern of cyclin E in tumour cells. Int J Cancer. 2003 Apr 10;104(3):369-75. PMID: 12569561
9. Lindblad J, **Wählby C**, Bengtsson E, Zaltsman A. Image analysis for automatic segmentation of cytoplasms and classification of Rac1 activation. Cytometry A. 2004 Jan;57(1):22-33. PMID: 14699602
10. Jarvius M, Paulsson J, Weibrecht I, Leuchowius KJ, Andersson AC, **Wählby C**, Gullberg M, Botling J, Sjöblom T, Markova B, Ostman A, Landegren U, Söderberg O. In situ detection of phosphorylated platelet-derived growth factor receptor beta using a generalized proximity ligation method. Mol Cell Proteomics. 2007 Sep;6(9):1500-9. Epub 2007 Jun 12. PMID: 17565975
11. Jahangir Tafrechi RS, van de Rijke FM, Allallou A, Larsson C, Sloos WC, van de Sande M, **Wählby C**, Janssen GM, Raap AK. Single-cell A3243G mitochondrial DNA mutation load assays for segregation analysis. J Histochem Cytochem. 2007 Nov;55(11):1159-66. Epub 2007 Aug 6. PMID: 17679731
12. Allalou A, **Wählby C**. BlobFinder, a tool for fluorescence microscopy image cytometry. Comput Methods Programs Biomed. 2008 Oct 16. [Epub ahead of print] PMID: 18950895

13. Göransson J, **Wählby C**, Isaksson M, Howell WM, Jarvius J, Nilsson M. A single molecule array for digital targeted molecular analyses. Nucleic Acids Res. 2009 Jan;37(1):e7. Epub 2008 Nov 25. PMID: 19033366
14. Pinidiyaarachchi A, Zieba A, Allalou A, Pardali K, **Wählby C**. A detailed analysis of 3D subcellular signal localization. Cytometry A. 2009 Apr;75(4):319-28. PMID: 19006073
15. Zieba A, **Wählby C**, Hjelm F, Jordan L, Berg J, Landegren U, Pardali K. Bright-field microscopy visualization of proteins and protein complexes by in situ proximity ligation with peroxidase detection. Clin Chem. 2010 Jan;56(1):99-110. Epub 2009 Nov 19.PMID: 19926775

D. Research Support

Ongoing Research Support

Swedish Research Council Collaboration Grant, Medicine 01/01/08-12/31/10

A multidisciplinary approach to establish mechanisms for mitochondrial DNA segregation in human disease

In this project we apply a powerful fluorescent *in situ* hybridization (FISH) technology to follow transmission of heteroplasmic mtDNA mutations in real tissues *in situ* and develop image analysis software to do three-dimensional (3D) reconstructions of the distribution of mutated mtDNA molecules in mammalian tissues.

Role: co-Principal Investigator, with Mats Nilsson and Nils-Göran Larsson

Completed research support

EU-grant for SMEs in Health Research 08/01/06-07/31/09

Life sciences, genomics and biotechnology for health, ENLIGHT.

The aim of the project was to develop molecular methods and image analysis tools for analysis of cancer biomarkers *in situ*.

Role: Project manager, WP2: Image analysis

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Carpenter, Anne E.	Director, Imaging Platform at the Broad Institute of Harvard and MIT
eRA COMMONS USER NAME (credential, e.g., agency login) XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Purdue University, West Lafayette	B.S.	05/97	Biology
University of Illinois, Urbana-Champaign	Ph.D.	05/03	Cell Biology
Whitehead Institute for Biomed. Research, Cambridge MA	postdoc	12/06	Image analysis for HTS

A. Personal Statement

The primary work in this proposal is the development, validation, application, and dissemination of image processing algorithms for high-throughput images, which is the focus of my research group. Our work on *C. elegans,* specifically, began six years ago and constituted both guidance on sample preparation and image acquisition as well as image analysis methods. This work yielded the establishment of a *C. elegans* screening facility at MGH by Fred Ausubel and Gary Ruvkun (Publication #2 in section C) and the completion of the first high-throughput *C. elegans* screen analyzed by automated image analysis, which identified novel inhibitors of infection by the pathogen *E. faecalis* (Pub. #4). The image analysis approach we developed is now part of my group's software package, CellProfiler, the first open-source software designed for high-throughput microscopy (Pubs. #1,3,12,15). CellProfiler has met worldwide success, being cited by more than 150 papers within 3 years of being published. My group is now focused on extracting rich information from complex image-based screens with dozens of collaborators probing diverse biological questions in the Boston area and around the world, resulting in a fairly unconventional publication list, with a large number of high-impact collaborative papers (most not listed in Section C for lack of space). For example, 10 of the 17 papers from my 3.5 year postdoc were driven by applications of our methods to various biological areas of interest.

Based on this track record, and now that Dr. Wählby has established herself leading the *C. elegans* projects (Pub. #5), I will play a supporting role. Creating and applying software is my group's expertise and requires a well-integrated team; my role on the project will primarily be to guide the project team members that will work side by side with the *C. elegans* collaborators to produce and disseminate user-friendly software incorporating the algorithms developed by [redacted]

B. Positions and Honors

Positions and Employment:

1997-2003 PhD Student w/ Andrew S. Belmont, MD, PhD, University of Illinois at Urbana-Champaign
2003-2006 Postdoctoral Fellow w/ David Sabatini, MD, PhD, Whitehead Institute for Biomedical Research
2007-Present Director, Imaging Platform at the Broad Institute of Harvard and MIT

Other Experience and Professional Memberships

Professional memberships: Phi Kappa Phi (1997), Phi Beta Kappa (1997), Society for Biomolecular Screening (2002), Computational & Systems Biology Initiative of MIT (2003), International Society for Analytical Cytology (2005), American Society for Cell Biology (2005), Association for Women in Science (2006), American Association for the Advancement of Science (2006), Massachusetts Academy of Sciences (2008).

2008 Co-chair, BioImage Informatics Workshop, Santa Barbara, CA
2008 Ad hoc member, NIH Microscopy Imaging study section
2010 Co-organizer, Cold Spring Harbor Labs conf.: "Automated imaging and high-throughput phenotyping"

Honors

1994 National Merit Scholar, Presidential Scholar semifinalist
1995 Pew Research scholarship (declined)
1996 Howard Hughes Medical Institute undergraduate research fellowship
1997 Fellowships from Phi Kappa Phi, the University of Illinois, UI's NIH training grant
1998 Howard Hughes Medical Institute predoctoral fellowship (5 years)
1998 National Science Foundation fellowship (declined)
2003 BioVision World Leader delegate
2003 Merck/MIT Computational & Systems Biology Initiative fellowship
2004 Life Sciences Research Foundation fellowship (3 years)
2006 L'Oreal USA Women in Science fellowship
2007 Genome Technology "Rising young investigator"
2008 Featured in public television show "Bold Visions: Women in Science and Technology"
2008 Elected to be a fellow of the Massachusetts Academy of Sciences
2009 Awarded Bio-IT World Best Practices Award for CellProfiler software project

C. Selected Peer-reviewed publications (chosen from 36 since first paper in 2001)
Researchers contributing to the present proposal are underlined below

Most relevant to the current application

1. **Carpenter AE**, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology, 7:R100. PMID: 17076895 PMCID: PMC1794559
2. Vokes, MS, **Carpenter AE** (2008) Using CellProfiler for automatic identification and measurement of biological objects in images. In: Ausubel FM et al., eds. Current Protocols in Molecular Biology 82:14.17.1-14.17.12. PMID: 18425761
3. Jones TR*, **Carpenter AE*** (the first two authors contributed equally), Lamprecht MR, Moffat J, Silver S, Grenier J, Castoreno AB, Eggert US, Root DE, Golland P, Sabatini DM (2009) Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. PNAS 106(6):1826-1831/doi:10.1073/pnas.0808843106. PMID: 19188593 PMCID: PMC2634799
4. Moy TI, Conery AL, Larkins-Ford J, Wu G, Mazitschek R, Casadei G, Lewis K, **Carpenter AE**, Ausubel FM (2009) High throughput screen for novel antimicrobials using a whole animal infection model. ACS Chemical Biology 4/doi:10.1021/cb900084v. PMID: 19572548 PMCID: In process
5. Wählby C, Riklin-Raviv T, Ljosa V, Conery AL, Golland P, Ausubel FM, **Carpenter AE** (2010) Resolving clustered worms via probabilistic shape models. Accepted for publication in proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI) 2010

Additional recent publications of importance to the field (in chronological order)

6. **Nye, A Carpenter** (former name), Rajendran RR, Stenoien DL, Mancini MA, Katzenellenbogen BS, Belmont AS (2002) Alteration of large-scale chromatin structure by estrogen receptor. Molecular and Cellular Biology, 22(10):3437-49. PMID: 11971975 PMCID: PMC133805
7. **Carpenter AE**, Sabatini DM (2004) Systematic genome-wide screens of gene function. Nature Reviews Genetics, 5(1):11-22. PMID: 14708012
8. **Carpenter AE**, Memedula S, Plutz MJ, Belmont AS. Common effects of acidic activators on large-scale chromatin structure and transcription (2005) Molecular and Cellular Biology, 25(3):958-968. PMID: 15657424 PMCID: PMC544008
9. Jones TR, **Carpenter AE**, Golland P (2005) Voronoi-based segmentation of cells on image manifolds. Proceedings of the Workshop on Computer Vision for Biomedical Image Applications (CVBIA). Yanxi Liu, Tianzi Jiang, Changshui Zhang (Eds.). Beijing, China, October 21. Lecture Notes in Computer Science 3765. Published by Springer-Verlag, Berlin, p. 535-543, ISBN 3-540-29411-2
10. Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, **Carpenter AE**, Foo SY, Stewart SA, Stockwell BR, Hacohen N, Hahn WC, Lander ES, Sabatini DM, Root DE (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell, 124(6):1283-98. PMID: 16564017
11. Jones TR, **Carpenter AE**, Sabatini DM, Golland P (2006) Methods for high-content, high-throughput image-based cell screening. Proceedings of the Workshop on Microscopic Image Analysis with

Applications in Biology (MIAAB). Metaxas DN, Whitaker RT, Rittcher J, Sebastian T (Eds). Copenhagen, Denmark, October 5, pp 65-72.
12. Lamprecht MR, Sabatini DM, **Carpenter AE** (2007) CellProfiler: free, versatile software for automated biological image analysis. Biotechniques. 42(1):71-75.
13. **Carpenter AE** (2007) Extracting rich information from images. In: Clemons PA et al., eds. Cell-Based Assays for High-Throughput Screening, Methods in Molecular Biology 486:14. New York, NY: Humana Press; 193-211. PMID: 19347625
14. **Carpenter AE** (2007) Image-based chemical screening. Nature Chemical Biology 3:461-465. PMID: 17637778
15. Jones, TR, Kang IH, Wheeler DB, Lindquist RA, Papallo A, Sabatini DM, Golland P, **Carpenter AE** (2008) CellProfiler Analyst: data exploration and analysis software for complex image-based screens. BMC Bioinformatics 9(1):482/doi: 10.1186/1471-2105-9-482. PMID: 19014601 PMCID: PMC261443

D. Research Support

Ongoing Research Support

AstraZeneca Pharmaceuticals Collaboration (PI: Anne Carpenter) 09/08/08 - 10/31/10
Image Data Mining for Determining Small Molecule Mechanisms of Action
In this project, we are developing methods to analyze images from four image-based screens, provided by AstraZeneca. The hundreds of morphological features extracted from these images are compared to data AstraZeneca has already extracted, using commercial software, in terms of their power to discriminate mechanisms of action among a small library of chemical compounds.
Role: PI

NIH R01-AI085581 (PI: Fred Ausubel) 09/28/09 - 08/31/14
Identifying novel anti-infectives by high throughput screening in whole animals
The goal of this project is to perform a screen of 250,000 chemical compounds to identify new classes of anti-infectives against *Pseudomonas aeruginosa.* Promising compounds will undergo characterization, efficacy testing in other gram-negative bacteria (Klebsiella, Acinetobacter, Enterobacter), testing in mouse models of infection, and in some cases molecular target identification.
Role: PI of Subaward

NIH RC2-OD-09-004 (PI: Kevin Eliceiri) 09/30/09 - 09/29/11
ImageJ as an extensible image processing framework
The major goals of this grant are to improve ImageJ's core architecture, expand its functionality by interfacing ImageJ with existing open-source programs, and grow community-driven development while maintaining compatibility.
Role: PI of Subaward

NIH U54-HG005032-01 (PI: Stuart Schreiber) 09/01/08 – 05/31/14
Broad Institute Comprehensive Screening Center MLPCN
The BCSC will provide a wide range of assay development, assay adaptation/implementation, high-throughput screening via automation, follow-up and medicinal chemistry, informatics, and project management through five core groups. The investigators with primary responsibilities for these functions have come together in an open data-sharing environment to form a single, integrated pipeline where the concepts of one activity or discipline affect the thinking of the other.
Role: Key personnel

NIH 5 RL1-CA133834-03 (PI: Todd Golub) 10/01/07 - 09/30/12
Genomics Based Drug Discovery
This project proposes to create a new approach to drug discovery involving four components: (1) Leadership; (2) a discovery pipeline; (3) target ID; and (4) testing on 6 driving medical projects.
Role: Key personnel

Completed Research Support

Broad Institute SPARC Proposal 08/01/06-02/28/09
Dramatically Enhancing the Throughput of *C.elegans* RNAi and Drug Screens
The overall goal of this project is to increase the throughput of *C. elegans* RNAi and chemical screens by 10 to 100 fold.
Role: participant (PIs: Ausubel and Ruvkun)

Eli Lilly Collaboration 04/01/09-12/31/09
Enabling phenotypic screens with multiple complex cellular morphologies
Role: PI

Culpeper Foundation Biomedical Pilot Grant 06/01/06-05/31/08
Testing for Drug Targets in Realistic Cell Environments. In this project, we explored a system to test multiple genes for involvement in disease by implanting RNAi microarrays in whole animals, including image analysis to quantify the resulting cellular phenotypes.
Role: PI

Merck collaboration 10/01/06-12/30/07
RNAi /Compound Screening in Drosophila for Target/Pathway ID and to Identify Proteins Involved in DNA Damage Repair. In this project, we screened Drosophila genes by RNAi cell microarrays and using advanced image analysis methods to identify appropriate drug targets.
Role: co-PI, with David M. Sabatini

BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Frederick M. Ausubel	Molecular Biologist, Massachusetts General Hospital Professor of Genetics, Harvard Medical School
eRA COMMONS USER NAME XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
University of Illinois, Urbana, IL	B.S.	1966	Chemistry
Massachusetts Institute of Technology, Cambridge, MA	Ph.D.	1972	Biology

A. Personal Statement

My research group will be responsible for completing the laboratory work for the *C. elegans* infection screens, which will provide images necessary for this project. We will also follow up on the small-molecule regulators of infection and immunity that are identified in the screens. My group has pioneered the development of so-called multi-host pathogenesis systems that involve the infection of invertebrate hosts with human bacterial pathogens. Together with Gary Ruvkun's laboratory, we have also pioneered and assembled an automated *C. elegans* sample preparation pipeline over the past three years to enable high-throughput image-based screens. We used the system to screen 37,000 unique compounds for regulators of infection of *C. elegans* by the important human opportunistic pathogen *Enterococcus faecalis*—to our knowledge, the first screen utilizing adult *C. elegans* animals to employ automated image analysis. This work was completed in collaboration with the Broad Institute's Imaging Platform, another collaborator on this proposal. Members of my group have been developing several new assays described in the proposal and have been working closely with both the Imaging Platform and Dr. Carolina Wählby in this effort.

B. Positions and Honors

EXPERIENCE/EMPLOYMENT:

1966 - 1971: Graduate Student, M.I.T. Purification and properties of bacteriophage lambda integrase.
1972 - 1973: Instructor and Research Associate, M.I.T. Genetic analysis of nitrogen fixation genes.
1974 - 1975: Research Fellow, Harvard University. Genetic analysis of nitrogen fixation genes.
1975 - 1982: Assistant and Associate Professor of Biology, Department of Cellular and Developmental Biology, Harvard University. Molecular-genetic analysis of nitrogen fixation genes.
1982 - Professor of Genetics, Harvard Medical School. Molecular genetics of microbial pathogenesis and host defense.

AWARDS:

Member, National Academy of Sciences, 1994
Member, American Academy of Microbiology, 2002
Member, American Academy of Arts and Sciences, 2003

C. 15 selected relevant publications from a total of 297

1. Mahajan-Miklos, S., M.-W. Tan, L.G. Rahme and F.M. Ausubel (1999) Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa-Caenorhabditis elegans* pathogenesis model. *Cell* 96:47-56. PMID: 9989496
2. Tan, M.-W., S. Mahajan-Miklos, and F.M. Ausubel (1999) Killing of *C. elegans* by *P. aeruginosa* used to model mammalian bacterial pathogenesis. *Proc. Natl. Acad. Sci. USA* 96:715-720. PMID: 9892699; PMCID: PMC15202
3. Aballay, A., P. Yorgey and F.M. Ausubel (2000) *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Current Biology* 10:1539-1542. PMID: 11114525

4. Kim, D.H., R. Feinbaum, G. Alloing, F.E. Emerson, D.A. Garsin, H. Inoue, M. Tanaka-Hino, N. Hisamoto, K. Matsumoto, M.-W. Tan and F.M. Ausubel (2002) A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* 297:623-626. PMCID: PMC128160
5. Garsin, D.A., J. Villanueva, J. Begun, C.D. Sifri, D.H. Kim, S.B. Calderwood, G.B. Ruvkun and F.M. Ausubel (2003) Long-lived *C. elegans daf-2* mutants are resistant to bacterial pathogens. *Science* 300:1921. PMID: 12817143
6. Liberati, N.T., K.A. Fitzgerald, D.H. Kim, R. Feinbaum, D.T. Golenbock, and F.M. Ausubel (2004) Requirement for a conserved Toll/interleukin-1 resistance domain protein in the *Caenorhabditis elegans* immmune response. *Proc. Natl. Acad. Sci. USA* 101:6593-6598. PMCID: PMC404090
7. Moy, T.I., A.R. Ball, Z. Anklesaria, G. Casadei, K. Lewis and F.M. Ausubel (2006). Identification of novel antimicrobials using a live-animal infection model. *Proc. Natl. Acad. Sci. USA* 103:10414-10419. PMCID: PMC1482800
8. Troemel, E.R., S.W. Chu, V. Reinke, S.S. Lee, F.M. Ausubel and D.H. Kim (2006) p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans. PLoS Genet.* 2:e183. PMCID: PMC1635533
9. Breger, J., G. Aperis, B.B. Fuchs, T.I. Moy, F.M. Ausubel and E. Mylonakis (2007) Antifungal chemical compounds identified using a *C. elegans* pathogenicity assay. *PLoS Pathog.* 3:e18. doi:10.1371/journal.ppat.0030018. PMCID: PMC1790726
10. Troemel E.R., M.-A. Felix, N.K. Whiteman, A. Barriere, and F.M. Ausubel (2008) Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. *PLoS Biol* 6:e309. PMCID: PMC2596862
11. Irazoqui, J.E., A. Ng, R.J. Xavier, and F.M. Ausubel (2008). Role for β-catenin and HOX transcription factors in *Caenorhabditis elegans* and mammalian host epithelial-pathogen interactions. *Proc. Natl. Acad. Sci. USA* 105:17469-17474. PMCID: PMC2582251
12. Powell, J.R., D.H. Kim and F.M. Ausubel (2009) The G protein-coupled receptor FSHR-1 is required for the *Caenorhabditis elegans* innate immune response. *Proc. Natl. Acad. Sci. USA*, 106:2782-2787. PMCID: PMC2650343.
13. Moy, T.I., A.L. Connery, J. Larkins-Ford, G. Wu, R. Mazitschek, G. Casadei, K. Lewis, A.E. Carpenter and F.M. Ausubel (2009) High throughput screen for antimicrobials using a whole animal infection model. *ACS Chem. Biol,* 4:527-533. PMCID: PMC2745594
14. Estes, K., T. Dunbar, J. Powell, F.M. Ausubel and E.R. Troemel (2010) The bZIP transcription factor zip-2 mediates an early response to *P. aeruginosa* infection in *C. elegans*. *Proc. Natl. Acad. Sci.* USA, in press.
15. Wählby, C., T. Riklin-Raviv, V. Ljosa, A.L. Conery, P. Golland, F.M. Ausubel and A.E. Carpenter (2010) Resolving clustered worms via probablistic shape models. Accepted for publication in proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI) 2010.

D. Selected Research Support (ongoing or completed in past 3 years) Relevant to Proposal

NIAID 1 R01 AI072508-03 (F.M. Ausubel, PI) 12/15/06-11/30/09
Novel whole-animal screens for anti-microbials
The goal of this project was to develop high throughput screens for identifying compounds that will cure a persistent bacterial infection in the intestine of the nematode worm *Caenorhabditis elegans.*

NIAID 1P01AI083214-01 (M. Gilmore, PI) (subcontract Schepens Eye Institute) 09/01/09 – 08/31/11
Harvard-wide program on antibiotic resistance
This is a multi-investigator program project grant. The title of the Ausubel subproject is: “Identification of pathways that can be targeted for the development of novel therapies for MRSA.” The goal is to use whole-animal high throughput assays to identify antimicrobial compounds that cure *C. elegans* of an MRSA infection.

NIAID 1 R01 AI076372-01A2 (K. Lewis, PI) (subcontract Northeastern University) 12/1/08-11/30/2013
A synergy-based therapy against *C. difficile*
This is a multi-investigator project. The aim of this project is to use a *C. elegans* infection model to identify low molecular weight compounds that function as MDR pump inhibitors that can be used in combination with a weak antibiotic to potentiate the activity of the antibiotic.

NIAID (Transformative) 1 R01 AI085581-01 (F.M. Ausubel, PI) 9/1/09-8/31/14
Identifying novel anti-infectives by high through-put screening in whole animals
The goal of this project is to use a high throughput screening assay involving whole *C. elegans* animals to identify low molecular weight compounds that cure *C. elegans* of a *Pseudomonas aeruginosa* infection.

BIOGRAPHICAL SKETCH

NAME	POSITION TITLE
Golland, Polina	Associate Professor of Computer Science and Electrical Engineering
eRA COMMONS USER NAME (credential, e.g., agency login) XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Technion, Haifa, Israel	B.A.	05/93	Computer Science
Technion, Haifa, Israel	M.Sc.	05/95	Computer Science
Massachusetts Institute of Technology, Cambridge, MA	Ph.D.	05/01	Electrical Engineering and Computer Science (Biomedical Image Analysis)
Massachusetts Institute of Technology, Cambridge, MA	Postdoc	10/03	Biomedical Image Analysis

A. Personal Statement

My expertise is in biomedical image analysis. My group has demonstrated novel methods for image segmentation and registration, shape analysis and fMRI modeling. In particular, we have developed novel methods for shape analysis that enable us to capture shape variation within a population and techniques for image segmentation that utilize prior information, such as shape models, to achieve accurate delineation of objects of interest in the images. Our collaboration with the Imaging Platform at the Broad Institute dates back several years. We jointly developed methods for cell segmentation and analysis that represent the core of CellProfiler, a software platform used by many biology groups for high throughput analysis of microcopy images.

I will contribute to the proposed project by collaborating with Dr. Wählby and her team on developing methods for worm segmentation in microscopy images. Specifically, we will apply shape modeling techniques to worm segmentation, further refining the approach demonstrated in the preliminary studies. In addition, we will actively collaborate with Dr. Wählby's group on building population models based on shape and appearance of the worms, to enable phenotype scoring in high throughput experiments. Dr. Riklin-Raviv is currently a postdoc in my group; her expertise is in shape-based segmentation. She will develop proposed methods under my guidance and in direct collaboration with Dr. Wählby.

B. Positions and Honors

Positions and Employment

2001-2003 Postdoctoral Research Associate, Artificial Intelligence Laboratory, MIT, Cambridge, MA.
2003-2008 Assistant Professor, EECS Department, MIT, Cambridge, MA.
2008-present Associate Professor, EECS Department, MIT, Cambridge, MA.
2009-present Associate Member of the Broad Institute of Harvard and MIT, Cambridge, MA.

Other Experience and Professional Memberships

Member of the Board of Directors, International Society on Medical Image Computing and Computer-Assisted Intervention.
Associate Editor, IEEE Transactions on Medical Imaging (TMI).
Member of the Editorial Board, Journal of Medical Image Analysis (MedIA).
Member of the Editorial Board, NeuroImage.
NIH Panel on Biomedical Imaging, member (2007), ad-hoc reviewer (2009).

NSF Panel on Collaborative Research in Computational Neuroscience (CRCNS), member (2006,2008), ad-hoc reviewer (2007).

Honors

1998 Marr Prize, Honorable Mention for paper "Stereo Matching with Transparency and Matting", ICCV: Sixth International Conference On Computer Vision.

1999 Best Poster Award, Honorable Mention for paper "Statistical Shape Analysis Using Fixed Topology Skeletons: Corpus Callosum Study", IPMI: International Conference on Information Processing and Medical Imaging.

2007 NSF CAREER Award. “Computational Modeling of Spatial Activation Patterns in fMRI.”

2007 MICCAI Young Scientist Award, for paper “Effects of Registration Regularization and Atlas Sharpness on Segmentation Accuracy”, B.T.T. Yeo, M.R. Sabuncu, R. Desikan, B. Fischl, P. Golland. MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention.

2007-10 Distinguished Alumnus (class of ‘64) Career Development Chair.

2009 MICCAI Young Scientist Award, for paper “Joint Segmentation of Image Ensembles via Latent Atlases”, T. Riklin Raviv, K. Van Leemput, W.M. Wells III, P. Golland. MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention.

2009 Erbsmann Award, Honorable Mention, for paper “Exploratory fMRI Analysis without Spatial Normalization”, D. Lashkari and P. Golland. IPMI: International Conference on Information Processing and Medical Imaging.

C. Selected Peer-reviewed Publications

Most relevant to the current application

T. R. Jones, A. E. Carpenter, D. M. Sabatini, and P. Golland. Methods for High-Content, High-Throughput Image-Based Cell Screening. In Proc. the First MICCAI Workshop on Microscopic Image Analysis with Applications in Biology, 65-72, 2006.

A.E. Carpenter, T.R. Jones, M.R. Lamprecht, C. Clarke, I.H. Kang, O. Friman, D.A. Guertin, J.H. Chang, R.A. Lindquist, J. Moffat, P. Golland and D.M. Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology 7(10):R100, 2006.

T.R. Jones, I.-H. Kang, D.B. Wheeler, R.A. Lindquist, A. Papallo, D.M. Sabatini, P. Golland and A.E. Carpenter. CellProfiler Analyst: data exploration and analysis software for complex image-based screens. BMC Bioinformatics 9:482, 2008.

T.R. Jones, A.E. Carpenter, M.R. Lamprecht, J. Moffat, S.J. Silver, J.K. Grenier, A.B. Castoreno, U.S. Eggert, D.E. Root, P. Golland, and D.M. Sabatini. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. PNAS, A 106(6):1826-1831, 2009.

T. Riklin Raviv, K. Van Leemput, W.M. Wells III, and P. Golland. Joint Segmentation of Image Ensembles via Latent Atlases. In Proc. MICCAI: International Conference on Medical Image Computing and Computer Assisted Intervention, LNCS 5761:272-280, 2009. MICCAI Young Scientist Award.

Additional recent publications of importance to the field

P. Golland, W.E.L. Grimson, M.E. Shenton, R. Kikinis. Detection and Analysis of Statistical Differences in Anatomical Shape. Medical Image Analysis, 9(1):69-86, 2005.

P. Yu, P.E. Grant, Y. Qi, X. Han, F. Segonne, R. Pienaar, E. Busa, J. Pacheco, N. Makris, R.L. Buckner, P. Golland, B. Fischl. Cortical Surface Shape Analysis Based on Spherical Wavelets. IEEE Transactions on Medical Imaigng, 26(4):582-597, 2007.

B.T.T. Yeo, M.R. Sabuncu, R. Desikan, B. Fischl, and P. Golland. Effects of registration regularization and atlas sharpness on segmentation accuracy. Medical Image Analysis, 12(5):603-615, 2008.

B.T.T. Yeo, W. Ou and P. Golland. On the Construction of Invertible Filter Banks on the 2-Sphere. IEEE Transactions on Image Processing, 17(3):283-300, 2008.

Y. Golland, P. Golland, S. Bentin, and R. Malach. Data-driven clustering reveals a fundamental subdivision of the human cortex into two global systems. Neuropsychologia, 46(2):540-553, 2008.

D. Lashkari and P. Golland. Convex Clustering with Exemplar-Based Models. Advances in Neural Information Processing Systems, 20:825-832, 2008.

M.R. Sabuncu, S.K. Balci, M.E. Shenton, and P. Golland. Image-Driven Population Analysis Through Mixture Modeling. IEEE Transactions on Medical Imaging, 28(9):1473 - 1487, 2009.

W. Ou, M.S. Hamäläinen, and P. Golland. A Distributed Spatio-Temporal EEG/MEG Inverse Solver. NeuroImage, 44(3):932-946, 2009.

K. Van Leemput, A. Bakkour, T. Benner, G. Wiggins, L.L. Wald, J. Augustinack, B.C. Dickerson, P. Golland, and B. Fischl. Automated Segmentation of Hippocampal Subfields from Ultra-High Resolution In Vivo MRI. Hippocampus, 19:549-557, 2009.

B.T.T. Yeo, M.R. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl, and P. Golland. Spherical Demons: Fast Diffeomorphic Landmark-Free Surface Registration. IEEE Transactions on Medical Imaging, in press, 2009.

D. Research Support

Ongoing Research Support

NIH NAMIC U54-EB005149 (Kikinis PI) 11/1/04 – 7/31/10
National Alliance for Medical Image Analysis
The goal of this project is to create a national network of research centers focused on the creation of integrated tools for medical image analysis. Primary responsibility is development of novel algorithms for image segmentation and registration, and integration of those algorithms into a coordinated framework that incorporates tools from other sites.
Role: MIT site PI

NIH 1-R01-NS051826 (Grimson PI) 2/15/05-1/31/11
Computational Modeling of Anatomical Shape Distributions
This project aims to further develop and validate computational methods for representing anatomical shape and its variation in populations.
Role: Investigator

NSF IIS- 0642971 (Golland PI) 2/1/07-1/31/12
Computational Modeling of Spatial Activation Patterns in fMRI
This project aims to develop and deploy novel representations of spatial networks of co-activation in fMRI.
Role: PI

NIH NAC P41-RR13218 (Kikinis PI) 8/1/08 – 5/31/13
Neuroimaging Analysis Center (NAC)
The goal of this project is to develop fMRI analysis methods for characterizing functional connectivity patterns in normal subjects and clinical populations.
Role: MIT site PI

NSF IIS/CRCNS 0904625 (Golland PI) 9/01/09-8/31/12
Finding Structure in the Space of Activation Profiles in fMRI
This project aims to develop novel representations for brain organization based on fMRI experiments with rich stimulus sets and to apply the representations and the related algorithms to better characterize the ventral visual pathway using visual fMRI studies.
Role: PI

Completed Research Support

NIH mBIRN U24-RR021382 (Rosen PI) 10/1/04-5/31/09
Morphometry Biomedical Informatics Research Network (MBIRN)
The goals of this project are to develop a distributed computing platform for researchers in the field of computational anatomy. The particular responsibilities are to make the tools of machine learning and semantic information retrieval available to the computational anatomy community by incorporating them into the common analysis platform of mBIRN.
Role: Investigator

MINT (MIT internal grant) (Kanwisher, Golland PIs) 9/1/08 – 1/31/10
Discovering Structure in the Space of Activation Profiles
This project aims to develop novel representations for brain organization based on fMRI experiments with rich stimulus sets and to apply the representations and the related algorithms to better characterize the ventral visual pathway using visual fMRI studies.
Role: Co-PI

BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Gary Ruvkun	Molecular Biologist, MGH and Professor of Genetics, Harvard Medical School
eRA COMMONS USER NAME XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	YEAR(s)	FIELD OF STUDY
University of California at Berkeley	A.B.	1973	Biophysics
Harvard University	Ph.D.	1982	Biophysics

A. Personal Statement

My lab studies genetic pathways in the regulation of miRNA, siRNA, fat storage and aging. My laboratory has conducted numerous screens using genome-scale RNAi libraries for *C. elegans*, uncovering dozens of important genes with implications for human health that have been further investigated in my laboratory and my former postdocs' laboratories. Together with Fred Ausubel, I have worked to create an automated high-throughput *C. elegans* screening center, using fluorescence microscopy as the primary readout given the importance of visual phenotypes in the analysis of complex biological processes like metabolism and aging. The one remaining hurdle is image analysis. We have been collaborating with the Imaging Platform for three years and with Dr. Wahlby for the past year to overcome this challenge. We will make all of the images we have produced available for algorithm development, and we will use funding from other sources to complete assay development and genome-scale RNAi screening for regulators of fat accumulation and metabolism using oil red O staining, as well as the laboratory work to follow up on hits from this screen. We plan to contribute to the proposal by annotating control images for the oil red O assay, transferring images from the full-scale screen to the Wahlby group, and most importantly, to participate interactively in the development of algorithms for this assay by helping to assess the accuracy of automated scoring. Beyond the scope of the proposed grant to complete this particular screen, my laboratory has dozens of other large-scale screens that will make use of the algorithms and software developed, and we therefore are most supportive of continuing the collaboration.

B. Positions and Honors

1997-present Professor of Genetics, Harvard Medical School

Other Professional Activities:

1995 - present Editor, Developmental Biology
2004 - 2007 NIH National Advisory Council on Aging

Awards and Honors:

2002-present NIH Merit Award
2005 Rosenstiel Award from Brandeis University shared with Victor Ambros, Andy Fire, and Craig Mello.
2008 Benjamin Franklin Medal, Franklin Institute, with Victor Ambros and David Baulcombe
2008 Gairdner International Award, with Victor Ambros.
2008 National Academy of Sciences
2008 Lasker Basic Medical Research Award, Lasker Foundation, with Victor Ambros and David Baulcombe
2008 Warren Triennial Prize, Massachusetts General Hospital, with Victor Ambros
2009 Louisa Gross Horwitz Prize, Columbia University, with Victor Ambros
2009 American Academy of Arts and Sciences
2009 Massry Prize, with Victor Ambros

C. Selected peer reviewed publications (in chronological order) 131 publications total:

1. Curran, S. P. and G. Ruvkun. Lifespan regulation by evolutionarily conserved genes essential for viability. 2007. PloS Genetics, 2007 Apr 6;3(4):e56. Epub 2007 Feb 27. PMCID: PMC1847696
2. Samuelson, A. V., Carr, C. E., and G. Ruvkun. 2007. Gene activities that mediate increased lifespan of *C. elegans* insulin-like signaling mutants. Genes and Development, Nov 15;21(22):2976-94. PMCID: PMC2049198
3. Parry, D.H., Xu, J. and G. Ruvkun. 2007. A whole-genome RNAi screen for *C. elegans* miRNA pathway genes. Curr Biol. 2007 Dec 4;17(23):2013-22. Epub 2007 Nov 20. PMID: 18023351
4. Samuelson, A.V., R. R. Klimczak, D. Thompson, C. E. Carr, and G. Ruvkun. 2008 Identification of *C. elegans* genes regulating longevity using enhanced RNAi-sensitive strains. Cold Spring Harbor Symp Quant Biol.; 73: Circadian Rhythms. PMID: 18419309
5. Gabel, H. W. and G. Ruvkun. 2008. The exonuclease ERI-1 has a conserved dual role in 5.8S rRNA processing and RNAi. Nat Struct Mol Biol. 2008 May;15(5):531-3. Epub 2008 Apr 27. PMID: 18438419
6. Fischer, S.E.J., M. D. Butler, Q. Pan and G. Ruvkun. Trans-splicing in C. elegans generates the negative RNAi regulator ERI-6/7. Nature. 2008 Sep 25;455(7212):491-6. Epub 2008 Sep 10. PMID: 1878465
7. Wang, M. C., E. O'Rourke and G. Ruvkun. Fat metabolism links germline stem cells and longevity in C. elegans.. Science. 2008 Nov 7;322(5903):957-60. PMID: 18988854
8. Soukas, A. A., E. A. Kane, C. E. Carr, J. A. Melo, and G. Ruvkun. 2009. Rictor/TORC2 regulates fat metabolism, feeding, growth, and lifespan in *Caenorhabditis elegans.* Genes Dev. 2009 Feb 15;23(4):496-511. PMC2648650
9. Curran SP, Wu X, Riedel CG, Ruvkun G. 2009. A soma-to-germline transformation in long-lived Caenorhabditis elegans mutants. Nature 459: 1079-84. PMID: 19506556
10. O'Rourke EJ, Soukas AA, Carr CE, Ruvkun G. 2009. C. elegans major fats are stored in vesicles distinct from lysosome-related organelles. Cell Metab. 2009 Nov;10(5):430-5. PMID: 19883620

D. Research Support

Ongoing Research Support

PI: Gary Ruvkun:

5 R01 AG16636-11 ARRA 04/01/99-08/31/11
NIH Genetic and molecular basis of longevity
Aims 1.1. Classify stress responses activated in long-lived *C. elegans* using stress gene GFP fusion genes.
1.2 Endocrine exploration of the growth arrest pathways
2.1. Analyze the association of essential gene inactivation induced developmental arrest and longevity extension.
2.2. Identify components of the *cct-6* chaperonin complex stress response/soma to germline pathway activated in long-lived worms.
2.3. Molecular analysis of genes identified in the screens
3. Validation of *C. elegans* genes that act as co factors for *daf-2*-mediated increased longevity

5 R37 AG14161-14 05/01/96-08/31/11
NIH Inositol signaling in *C. elegans* senescence and diapause
Aim 1. Phenotypic characterization of gene inactivations and mutations in genes identified by RNAi that confer long lifespan
Aim 2. Identification of genes regulating longevity using enhanced RNAi-sensitive (ERI) strains.
Aim 3. Pathway analysis of new genes identified.

2 R01 GM044619-18 05/01/91-06/30/12
NIH Control of *C. elegans* lineage by heterochronic genes
Aim 1. Validation and classification of *C. elegans* genes that act as co factors for miRNA function
Aim 2. Pathway analysis of *C. elegans* genes that act as co factors for RNA interference

2 R01 DK070147-05 09/15/04-08/31/13
NIH Genetic and functional genomic analysis of *C. elegans* fat regulatory pathways
Aim 1: Metabolic and satiety profiling of serotonin signaling mutants in fat deposition
Aim 2: Molecular analysis of genes identified in the screens

NNG05GK27G 06/01/05-05/31/10
National Aeronautics and Space Administration, NASA
A Search for Extraterrestrial Genomes

NNX08AX15G 09/01/08-08/31/12
MIT/National Aeronautics and Space Administration, NASA
A Search for Extraterrestrial Genomes (SETG): An In-situ Detector for Life on Mars Ancestrally Related to Life on Earth

NNH08ZDA001N-MMAMA 06/01/09-12/01/10
National Aeronautics and Space Administration, NASA
REDGENES: Remotely Examining DNA with a Genetic Explorer in a Natural Extreme System

Glenn Foundation 03/01/09-03/15/11
Award for Research in Biological Mechanisms of Aging

3 R01 DK070147-06S1 ARRA 12/21/09-02/28/10
NIH Genomic analysis of *C. elegans* fat regulatory pathways

Completed Research Support

PI: Gary Ruvkun:

Broad Institute 08/01/06-02/28/09
SPARC Proposal
Dramatically Enhancing the Throughput of *C.elegans* RNAi and Drug Screens
Aims: The overall goal of this project is to increase the throughput of *C. elegans* RNAi and chemical screens by 10 to 100 fold.

BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

NAME	POSITION TITLE
Riklin-Raviv, Tamar	Postdoc researcher Computer Science / Artificial Intelligence Laboratory EECS department at MIT
eRA COMMONS USER NAME (credential, e.g., agency login) XXXXXXX	

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Hebrew University of Jerusalem, Israel	B.Sc.	09/93	Physics
Hebrew University of Jerusalem, Israel	M.Sc.	12/99	Computer vision
Tel-Aviv University, Israel	Ph.D.	02/08	Computer vision
Massachusetts Institute of Technology	Postdoc	current	Medical Image Analysis

A. Personal Statement

An essential component in the analysis of high-throughput images of *C. elegans* is the extraction of individual worms from the images; many infection- and metabolism-related phenotypes in the proposal depend on this. The almost identical appearance of the worms, due to their stereotypical development pattern, makes their shape an ideal cue for facilitating the segmentation. In the past few years I have been addressing problems of object segmentation in the presence of morphological information. In particular I developed variational methods based on an active contour paradigm known as level-set that allows the incorporation of implicit shape information within a unified cost functional for object segmentation. These approaches have garnered attention in the field; I have published 3 top-journal papers and 7 peer reviewed international conference papers (I am a first author in most of them) on this particular research direction and related topics. One of these won the prestigious young scientist award at a leading conference in the field of biomedical image analysis, the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI).

Since January 2008, I have pursued and expanded these research interests as a postdoctoral associate in the Medical Vision group of M.I.T., led by Prof. Polina Golland. I started collaborating with Dr. Wählby of the Broad Institute in the spring of 2009, continuing a long-term collaboration between Dr. Golland's group and the Broad.

Already, this work has been fruitful; our recent research on the segmentation of *C. elegans* in high-throughput images has been accepted for publication in the proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), and we have submitted a second paper together. Given my computational background in addition to my experience with *C. elegans* images so far, I am thus well-suited to continue working together with Dr. Wählby on developing novel algorithms for this important biological area.

B. Positions and Honors

Positions and Employment:

2003-2007 PhD student w/ Nahum Kiryati and Nir Sochen, School of Electrical Engineering, Tel-Aviv University
2008-present Postdoctoral associate w/ Polina Golland, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

Other Experience and Professional Memberships

Member, IEEE
2006- 2007 Coordinator of the Mathematical Visual Perception seminar at Tel-Aviv University
Ad-hoc journal reviewer: IEEE Transactions on Pattern Analysis and Machine Intelligence
IEEE Transactions on Signal Processing
IEEE Transactions on Systems, Man and Cybernetics
Program Committee: IEEE Conference on Computer Vision and Pattern Recognition 2006,2007,2009
IEEE International Conference on Computer Vision 2007

European Conference in Computer Vision 2004, 2006, 2008
International Journal of Biomedical Imaging
International Journal of Image and Graphics
Journal of Visual Communication and Image Representation

Honors

1993 The Hebrew University of Jerusalem Dean's list
2005 Weinstein award for excellent paper, "Unlevel-Sets: Geometry and Prior-based Segmentation"
2007 Fulbright Post-Doctoral fellowship
2007 The Commercial & Industrial Club Illan Ramon Post-doctoral Scholar
2007 Weinstein award for excellence in studies
2008 Weinstein award for excellent paper, "Prior-based Segmentation and Shape Registration in the Presence of Perspective Distortion"
2009 MICCAI 2009 young scientist award for "Joint Segmentation of Image Ensembles via Latent Atlases"

C. Peer-reviewed publications

Total number of citations according to Google scholar as of January 2010 is over 450.

Journal papers – reverse chronological order

1. T. Riklin-Raviv, N. Sochen and N. Kiryati, On Symmetry, Perspectivity and Level-set based segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI). Vol 31(8) pp 1458-1471, August 2009
2. T. Riklin-Raviv, N. Sochen and N. Kiryati, Shape based Mutual Segmentation. International Journal of computer Vision (IJCV). Vol 79(3) pp 231-245, September 2008
3. T. Riklin-Raviv, N. Kiryati and N. Sochen, Prior-based Segmentation and Shape Registration in the Presence of Perspective Distortion. International Journal of Computer Vision (IJCV). Vol 72(3) pp 309-328 May 2007
4. A. Shashua and T. Rikiln-Raviv, The Quotient Image: Class Based Re-Rendering and Recognition With Varying Illuminations. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI). Vol. 23(2) pp 129-139, February 2001.

Peer reviewed conference papers – reverse chronological order

5. C. Wählby, T. Riklin-Raviv, V. Ljosa, A.L. Conery, P. Golland, F.M. Ausubel, and A.E. Carpenter, Resolving Clustered Worms via Probabilistic Shape Model, IEEE International Symposium on Biomedical Imaging: From Nano to Micro (ISBI), April 2010, accepted for publication.
6. T. Riklin Raviv, K. Van-Leemput, W.M. Wells III and Polina Golland, Joint Segmentation of Image Ensembles via Latent Atlases, International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Part I, LNCS 5761, pp. 272–280, September 2009.

   Received the MICCAI 09 Young Scientist Award.
7. T. Riklin Raviv, B.M. Menze, K. Van-Leemput, B. Stieltjes, M.A. Weber, N. Ayache, W.M. Wells III and Polina Golland, Joint Segmentation via Patient-Specific Latent Anatomy Model, MICCAI workshop: Probabilistic Models for Medical Imaging Analysis (PMMIA), September 2009.
8. T. Riklin Raviv, N. Ben-Zadok and N. Kiryati Interactive Level-set Segmentation for Image Guided Therapy . IEEE International Symposium on Biomedical Imaging: From Nano to Micro (ISBI), pp. 1079-1082, June 2009.
9. N. Kiryati, T. Riklin Raviv, Y. Ivanchenko and S. Rochel, Real-time Abnormal Motion Detection in Surveillance Video. International Conference on Pattern Recognition (ICPR), pp. 1-4, December 2008.

10. T. Riklin-Raviv, N. Sochen, N. Kiryati, N. Ben-Zadok, S. Gefen, L. Bertand and J. Nissanov, Propagating Distributions for Segmentation of Brain Atlas. IEEE International Symposium on Biomedical Imaging: From Nano to Micro (ISBI), pp 1304-1307, April 2007.
11. T. Riklin-Raviv, N. Kiryati and N. Sochen, Segmentation with Level Sets and Symmetry. In Proc. of IEEE Conference on Computer Vision and Pattern Recognition. (CVPR), pp 1015-1022, June 2006.
12. T. Riklin-Raviv, N. Sochen and N. Kiryati, Mutual Segmentation with Level Sets. In the 5th IEEE Workshop on Perceptual Organization in Computer Vision (POCV) in conjunction with the CVPR. 2006.
13. T. Riklin-Raviv, N. Kiryati and N. Sochen, Prior-based Segmentation by Projective Registration and Level Sets. In Proc. of the Tenth IEEE International Conference on Computer Vision (ICCV).pp 204-211, October 2005.
14. T. Riklin-Raviv, N. Kiryati and N. Sochen, Unlevel-Sets: Geometry and Prior-based Segmentation. In Proc. of the European Conference on Computer Vision (ECCV). pp 50-61, May 2004.
15. T. Riklin-Raviv and A. Shashua, The Quotient Image: Class Based Recognition and Synthesis Under Varying Illumination Conditions. In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp 566-571, June 1999.

D. Research Support

none

PHS 398 Cover Page Supplement

OMB Number: 0925-0001

1. Project Director / Principal Investigator (PD/PI)

Prefix:
* First Name: Carolina
Middle Name:
* Last Name: Wahlby
Suffix: Ph.D.

2. Human Subjects

Clinical Trial? ☒ No ☐ Yes

* Agency-Defined Phase III Clinical Trial? ☐ No ☐ Yes

3. Applicant Organization Contact

Person to be contacted on matters involving this application

Prefix:
* First Name: Stacey
Middle Name:
* Last Name: Donnelly
Suffix:

* Phone Number: 617-714-7130
Fax Number: 617-714-8972
Email: awards@broadinstitute.org

* Title: Director, Sponsored Research and Planning

* Street1: 7 Cambridge Center
Street2:
* City: Cambridge
County/Parish:
* State: MA: Massachusetts
Province:
* Country: USA: UNITED STATES
* Zip / Postal Code: 02142-1401

PHS 398 Cover Page Supplement

4. Human Embryonic Stem Cells

* Does the proposed project involve human embryonic stem cells? ☒ No ☐ Yes

If the proposed project involves human embryonic stem cells, list below the registration number of the specific cell line(s) from the following list: http://stemcells.nih.gov/research/registry/. Or, if a specific stem cell line cannot be referenced at this time, please check the box indicating that one from the registry will be used:

Cell Line(s): ☐ Specific stem cell line cannot be referenced at this time. One from the registry will be used.

Tracking Number:GRANT10517004 Funding Opportunity Number:PA-10-067 Received Date:2010-02-03T16:41:19-04:00

PHS 398 Modular Budget, Periods 1 and 2

OMB Number: 0925-0001

Budget Period: 1

Start Date: 12/01/2010 End Date: 11/30/2011

A. Direct Costs

	* Funds Requested ($)
* Direct Cost less Consortium F&A	250,000.00
Consortium F&A	
* Total Direct Costs	250,000.00

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	* Funds Requested ($)
1.	Modified Total Direct Costs	64.4	202,480.00	130,397.00
2.				
3.				
4.				

Cognizant Agency (Agency Name, POC Name and Phone Number) DHHS, Louis Martillotti, 212-264-2069

Indirect Cost Rate Agreement Date 04/03/2009

Total Indirect Costs 130,397.00

C. Total Direct and Indirect Costs (A + B)

Funds Requested ($) 380,397.00

Budget Period: 2

Start Date: 12/01/2011 End Date: 11/30/2012

A. Direct Costs

	* Funds Requested ($)
* Direct Cost less Consortium F&A	250,000.00
Consortium F&A	
* Total Direct Costs	250,000.00

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	* Funds Requested ($)
1.	Modified Total Direct Costs	64.4	201,054.00	129,479.00
2.				
3.				
4.				

Cognizant Agency (Agency Name, POC Name and Phone Number) DHHS, Louis Martillotti, 212-264-2069

Indirect Cost Rate Agreement Date 04/03/2009

Total Indirect Costs 129,479.00

C. Total Direct and Indirect Costs (A + B)

Funds Requested ($) 379,479.00

PHS 398 Modular Budget, Periods 3 and 4

Budget Period: 3

Start Date: 12/01/2012 End Date: 11/30/2013

A. Direct Costs

	* Funds Requested ($)
* Direct Cost less Consortium F&A	250,000.00
Consortium F&A	
* Total Direct Costs	250,000.00

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	* Funds Requested ($)
1.	Modified Total Direct Costs	64.4	199,586.00	128,533.00
2.				
3.				
4.				

Cognizant Agency (Agency Name, POC Name and Phone Number): DHHS, Louis Martillotti, 212-264-2069

Indirect Cost Rate Agreement Date: 04/03/2009

Total Indirect Costs: 128,533.00

C. Total Direct and Indirect Costs (A + B) Funds Requested ($): 378,533.00

Budget Period: 4

Start Date: 12/01/2013 End Date: 11/30/2015

A. Direct Costs

	* Funds Requested ($)
* Direct Cost less Consortium F&A	250,000.00
Consortium F&A	
* Total Direct Costs	250,000.00

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	* Funds Requested ($)
1.	Modified Total Direct Costs	64.4	198,074.00	127,560.00
2.				
3.				
4.				

Cognizant Agency (Agency Name, POC Name and Phone Number): DHHS, Louis Martillotti, 212-264-2069

Indirect Cost Rate Agreement Date: 04/03/2009

Total Indirect Costs: 127,560.00

C. Total Direct and Indirect Costs (A + B) Funds Requested ($): 377,560.00

PHS 398 Modular Budget, Periods 5 and Cumulative

Budget Period: 5

Start Date: 12/01/2014 End Date: 11/30/2015

A. Direct Costs

	* Funds Requested ($)
* Direct Cost less Consortium F&A	250,000.00
Consortium F&A	
* Total Direct Costs	250,000.00

B. Indirect Costs

	Indirect Cost Type	Indirect Cost Rate (%)	Indirect Cost Base ($)	* Funds Requested ($)
1.	Modified Total Direct Costs	64.4	196,516.00	126,556.00
2.				
3.				
4.				

Cognizant Agency (Agency Name, POC Name and Phone Number) DHHS, Louis Martillotti, 212-264-2069

Indirect Cost Rate Agreement Date 04/03/2009

Total Indirect Costs 126,556.00

C. Total Direct and Indirect Costs (A + B)

Funds Requested ($) 376,556.00

Cumulative Budget Information

1. Total Costs, Entire Project Period

*Section A, Total Direct Cost less Consortium F&A for Entire Project Period	$	1,250,000.00
Section A, Total Consortium F&A for Entire Project Period	$	
*Section A, Total Direct Costs for Entire Project Period	$	1,250,000.00
*Section B, Total Indirect Costs for Entire Project Period	$	642,525.00
*Section C, Total Direct and Indirect Costs (A+B) for Entire Project Period	$	1,892,525.00

2. Budget Justifications

Personnel Justification	1246-Personneljustification.pd	Add Attachment	Delete Attachment	View Attachment
Consortium Justification		Add Attachment	Delete Attachment	View Attachment
Additional Narrative Justification	1247-FAexplaination.pdf	Add Attachment	Delete Attachment	View Attachment

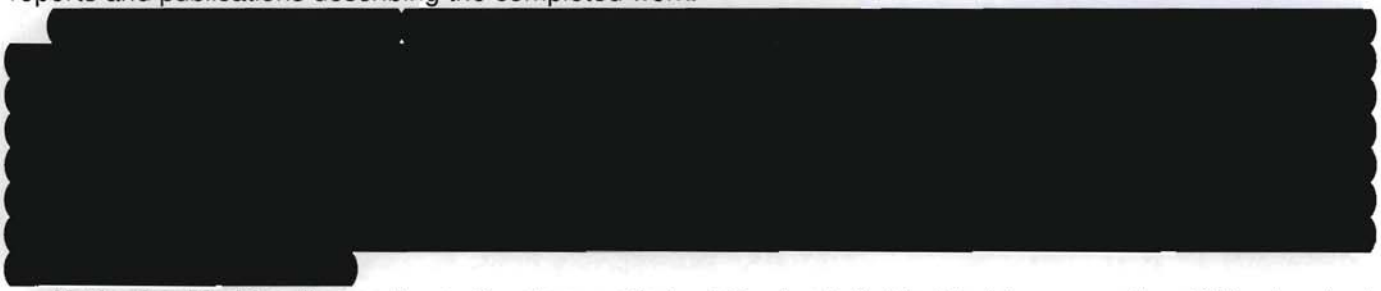
Tracking Number:GRANT10517004 Funding Opportunity Number:PA-10-067 Received Date:2010-02-03T16:41:19-04:00

Budget Justification

Personnel

Carolina Wahlby, Ph.D. (Principal Investigator, effort = 8.4 calendar). Dr. Wahlby is the principal investigator for the project and will have primary responsibility for developing and validating algorithms for *C. elegans*, in addition to coordinating the collaborators and contributors on the project. Dr. Wahlby earned her PhD in digital image processing and did her post-doctoral work at the Dept. of Genetics and Pathology at Uppsala University. Many of the algorithms she developed during her PhD work are used worldwide for high-throughput cell image analysis. Prior to joining the Broad Institute, she was promoted to Associate Professor at the Centre for Image Analysis, Uppsala University, Sweden, an appointment she still holds. Dr. Wahlby has supervised 4 graduate students in digital image processing, developing image-analysis algorithms for biomedical applications, and she has more than 10 years of experience in the field. She has managed a project of similar scale to the present proposal, for image-analysis development in an EU-financed cross-disciplinary project involving research groups and private companies in five different European countries. Since joining the Broad Institute, Dr. Wahlby began collaborations with the Ausubel, Carpenter, Golland, and Ruvkun groups, focused on developing algorithms for *C. elegans* high-throughput screening.

Anne Carpenter, Ph.D. (Imaging Platform Director, effort = 1 calendar). Dr. Carpenter leads a research group of computer scientists and biologists. Dr. Carpenter's team developed CellProfiler, the first open-source software package designed for high-throughput cell image analysis, as well as CellProfiler Analyst, software for the interactive exploration and analysis of the resulting multidimensional data. CellProfiler Analyst includes machine-learning–based scoring of complex and subtle cellular phenotypes. Her group has applied these award-winning tools to extract rich information from complex image-based screens with dozens of collaborators probing diverse biological questions in the Boston area and around the world. In recognition of this work, she was recently elected the youngest member of the Massachusetts Academy of Sciences. She worked with the Ausubel and Golland groups to successfully complete the first automated analysis of a high-throughput screen in adult *C. elegans*. Dr. Carpenter will manage the Imaging Platform team (described in Facilities and Resources) in their work to create, test, and disseminate high-throughput *C. elegans* software. She will also assist in writing reports and publications describing the completed work.



Katherine Madden (Image Assay Developer, effort = 4.8 calendar). Ms. Madden earned her BS in chemical engineering from the Massachusetts Institute of Technology. Since joining the Broad, she has developed many automated imaging assays, including two high-throughput screens comprised of hundreds of thousands of images, and several custom screens for subtle morphological phenotypes. As the primary image assay developer for the *C. elegans* projects since 2008, she has worked on assays to measure feeding on fluorescent beads, Nile Red accumulation in the gut, oil red O accumulation in fatty tissues, localization of an infection reporter, and expression of a marker for body muscle.

NOTES:

The F&A calculation is based on Broad's indirect cost rate of 64.4%. This project utilizes a Specialized Service Facility that is excluded from F&A.

OMB Number: 0925-0001

PHS 398 Research Plan

1. Application Type:

From SF 424 (R&R) Cover Page. The response provided on that page, regarding the type of application being submitted, is repeated for your reference, as you attach the appropriate sections of the Research Plan.

*Type of Application:

☒ New ☐ Resubmission ☐ Renewal ☐ Continuation ☐ Revision

2. Research Plan Attachments:

Please attach applicable sections of the research plan, below.

1. Introduction to Application (for RESUBMISSION or REVISION only)		Add Attachment	Delete Attachment	View Attachment
2. Specific Aims	1248-SpecificAims.pdf	Add Attachment	Delete Attachment	View Attachment
3. *Research Strategy	1249-ResearchStrategy.pdf	Add Attachment	Delete Attachment	View Attachment
4. Inclusion Enrollment Report		Add Attachment	Delete Attachment	View Attachment
5. Progress Report Publication List		Add Attachment	Delete Attachment	View Attachment

Human Subjects Sections

6. Protection of Human Subjects		Add Attachment	Delete Attachment	View Attachment
7. Inclusion of Women and Minorities		Add Attachment	Delete Attachment	View Attachment
8. Targeted/Planned Enrollment Table		Add Attachment	Delete Attachment	View Attachment
9. Inclusion of Children		Add Attachment	Delete Attachment	View Attachment

Other Research Plan Sections

10. Vertebrate Animals		Add Attachment	Delete Attachment	View Attachment
11. Select Agent Research		Add Attachment	Delete Attachment	View Attachment
12. Multiple PD/PI Leadership Plan		Add Attachment	Delete Attachment	View Attachment
13. Consortium/Contractual Arrangements		Add Attachment	Delete Attachment	View Attachment
14. Letters of Support	1250-LettersOfSupportCombine	Add Attachment	Delete Attachment	View Attachment
15. Resource Sharing Plan(s)	1251-ResourceSharingPlans.p	Add Attachment	Delete Attachment	View Attachment

16. Appendix Add Attachments | Remove Attachments | View Attachments

Specific Aims

Microscopy has emerged as one of the most powerful and informative ways to analyze cell-based high-throughput screening (HTS) samples in experiments designed to uncover novel drugs and drug targets. However, many diseases and biological pathways can be better studied in whole animals–particularly diseases that involve organ systems and multicellular interactions, such as metabolism and infection. The worm *Caenorhabditis elegans* is a well-established and effective model organism that can be robotically prepared and imaged, but existing image-analysis methods are insufficient for most assays.

We propose to develop algorithms for the analysis of high-throughput *C. elegans* images, validating them in three specific experiments to identify chemicals to cure human infections and genetic regulators of host response to pathogens and fat metabolism. Novel computational tools for automated image analysis of *C. elegans* assays will make whole-animal screening possible for a variety of biological questions not approachable by cell-based assays. Building on our expertise in developing image processing and machine learning algorithms for high-throughput screening, and on our established collaborations with leaders in *C. elegans* research, we will:

Aim 1: Develop algorithms for *C. elegans* viability assays to identify modulators of pathogen infection

Challenge: To identify individual worms in thousands of two-dimensional brightfield images of worm populations infected by Microsporidia, and measure viability based on worm body shape (live worms are curvy whereas dead worms are straight).

Approach: We will develop algorithms that use a probabilistic shape model of *C. elegans* learned from examples, enabling segmentation and body shape measurements even when worms touch or cross.

Impact: These algorithms will quantify a wide range of phenotypic descriptors detectable in individual worms, including body morphology as well as subtle variations in reporter signal levels.

Aim 2: Develop algorithms for *C. elegans* lipid assays to identify genes that regulate fat metabolism

Challenge: To detect worms versus background, despite artifacts from sample preparation, and detect subtle phenotypes of worm populations.

Approach: We will improve well edge detection, illumination correction, and detection of artifacts (e.g. bubbles and aggregates of bacteria) and enable image segmentation in highly variable image backgrounds using level-set segmentation. We will also design feature descriptors that can capture worm population phenotypes.

Impact: These algorithms will provide detection for a variety of phenotypes in worm populations. They will also improve data quality in other assays, such as those in Aims 1 and 3.

Aim 3: Develop algorithms for gene expression pattern assays to identify regulators of the response of the *C. elegans* host to *Staphylococcus aureus* infection

Challenge: To map each worm to a reference and quantify changes in fluorescence localization patterns.

Approach: We will develop worm mapping algorithms and combine them with anatomical maps to extract atlas-based measurements of staining patterns and localization. We will then use machine learning to distinguish morphological phenotypes of interest based on the extracted features.

Impact: These algorithms will enable addressing a variety of biological questions by measuring complex morphologies within individual worms.

In addition to discovering novel anti-infectives and genes involved in metabolism and pathogen resistance, this work will provide the *C. elegans* community with (a) a versatile, modular, open-source toolbox of algorithms readily usable by biologists to quantify a wide range of important high-throughput whole-organism assays, (b) a new framework for extracting morphological features from *C. elegans* populations for quantitative analysis of this organism, and (c) the capability to discover disease-related pathways, chemical probes, and drug targets in high-throughput screens relevant to a variety of diseases.

Primary collaborators

Gary Ruvkun and **Fred Ausubel**, MGH/Harvard Medical School: Development, execution, and follow-up of large-scale *C. elegans* screens probing metabolism and infection. **Polina Golland** and **Tammy Riklin-Raviv**, MIT Computer Science and Artificial Intelligence Lab: Illumination/bias correction, model-based segmentation, and statistical image analysis. **Anne Carpenter**, Broad Imaging Platform: Software engineering and support.

Research Strategy

A Significance

The NIH is committed to translating basic biomedical research into clinical practice and thereby impacting global human health[1], and Francis Collins identifies high-throughput technology as one of five areas of focus for the NIH's research agenda[2]. For many diseases, researchers have identified successful novel therapeutics or research probes by applying technical advances in automation to high-throughput screening (HTS) using either biochemical or cell-based assays[3–6]. Researchers are using genetic perturbations such as RNA interference or gene overexpression in cell-based HTS assays to identify genetic regulators of disease processes as potential drug targets[7–9]. However, the molecular mechanisms of many diseases that deeply impact human health worldwide are not well-understood and thus cannot yet be reduced to biochemical or cell-based assays.

Ideally, researchers could approach disease from a phenotypic direction, in addition to the traditional molecular approach, by searching for chemical or genetic regulators of disease processes in whole model organisms rather than isolated cells or proteins. Moving HTS towards more intact, physiological systems also improves the likelihood that the findings from such experiments accurately translate into the context of the human body (e.g., in terms of toxicity and bioavailability), simplifying the path to clinical trials and reducing the failure of potential therapeutics at later stages of testing. In fact, for some diseases, a whole organism screen may actually be necessary to break new therapeutic ground; in the search for novel therapeutics for infectious agents, for example, it is widely speculated that the traditional approach of screening for chemicals that directly kill bacteria *in vitro* has been largely exhausted[10]. Our work recently identified six novel classes of chemicals that cure model organisms from infection by the important human pathogen *E. faecalis* through mechanisms distinct from directly killing the bacterium itself[11]. Anti-infectives with new mechanisms of action are urgently needed to combat widespread antibiotic resistance in pathogens.

Enabling HTS in whole organisms is therefore recognized as a high priority (NIH PAR-08-024)[12,13]. *C. elegans* is a natural choice. Manually-analyzed RNAi and chemical screens are well-proven in this organism, with dozens completed[14–16]. Many existing assays can be adapted to HTS; instrumentation exists to handle and culture *C. elegans* in HTS-compatible multi-well. Its organ systems have high physiologic similarity and genetic conservation with humans[17,18]. *C. elegans* is particularly suited to assays involving visual phenotypes: physiologic abnormalities and fluorescent markers are easily observed because the worm is mostly transparent. The worms follow a stereotypic development pattern that yields identically-appearing adults[19,20], such that deviations from wild-type are more readily apparent.

The bottleneck that remains for tackling important human health problems using *C. elegans* HTS is image analysis (NIH PA-07-320)[21,22]. It has been recently stated, "Currently, one of the biggest technical limitations for large-scale RNAi-based screens in *C. elegans* is the lack of efficient high-throughput methods to quantitate lethality, growth rates, and other morphological phenotypes"[23]. **Our proposal to develop image analysis algorithms to identify regulators of infection and metabolism in high-throughput *C. elegans* assays would bring image-based HTS to whole organisms, and have the following impact:**

- **Identifying novel modulators of infection by the NIH priority pathogen Microsporidia** (Aim 1). Microsporidia are emerging human pathogens whose infection mechanisms are almost completely unknown. Further, they inflict agricultural damage and are on the EPA list of waterborne microbial contaminants of concern[24,25]. Identifying anti-microsporidian therapeutics is a special challenge because they are eukaryotes. Moreover, they are obligate intracellular pathogens so they are not amenable to traditional antibiotic screens; screening for drugs to kill them requires the presence of a validated, infectible host whose immune system is homologous to mammals, such as *C. elegans*[26,27]. This screen could identify not only useful chemical research probes and compounds that kill these pathogens outright, but also those that block microbial virulence, are modified by the host for full efficacy (prodrugs), or enhance host immunity.

- **Identifying novel regulators of fat metabolism** (Aim 2). Disregulation of metabolism results in many common and expensive chronic health conditions; diabetes alone affects 24 million Americans[28]. Energy centers must receive and integrate nutritional information from multiple peripheral signals across multiple tissues and cell types to elicit appropriate behavioral and metabolic responses; screening in a whole organism is important. In particular, screening with a strain of *C. elegans* with an RNAi-sensitive nervous system will likely reveal novel energy regulators of therapeutic and research value.

- **Identifying novel regulators of infection by the pathogen *Staphylococcus aureus*** (Aim 3). *S. aureus* is life-threatening for immune-compromised patients. Recently, antibiotic-resistant MRSA strains have created an urgent need for therapeutics with a new mechanism of action[29]. We will identify genetic regulators of the *C. elegans* host's response to infection by *S. aureus*[30]. These will lead to potential drug targets useful for boosting humans' innate immunity.

- **Enabling the automated analysis of a wide variety of *C. elegans* screens**. Because *C. elegans* has proven to be an excellent model for many human organs and processes, the impact of algorithms for automated scoring for currently intractable *C. elegans* image-based screens on our understanding and treatment of a variety of human diseases will be substantial. Adding novel *C. elegans* algorithms to existing open-source software will create a flexible toolbox that can be applied to other types of assays (including alternative formats such as microfluidics chambers; see Yanik support letter) with minimal modification:

  Aim 1: The algorithms developed for Aim 1 will enable scoring viability and other body morphology assays probing a number of biological processes. Our collaborators plan several RNAi and chemical screens using live/dead assays to identify modulators of many other clinically relevant pathogens (see Ausubel and Mylonakis support letters).

  Aim 2: The algorithms developed for the fat metabolism assay can also be used to quantify the levels of any stain within worms, to measure protein expression levels, the degree of staining by fluorescent dyes or antibodies, and promoter activity in reporter assays probing a wide range of biological processes.

  Aim 3: Where localization patterns are of interest, the algorithms developed for the gene expression pattern assay will often be directly applicable, especially given the proposed machine learning capabilities.

  Many benefits come from the automation of image analysis for such screens: (a) increased throughput so as to enable genome-scale RNAi and large-scale chemical screens in whole animals; (b) quantitative results amenable to data mining[31–33]; (c) increased objectivity and consistency; and (d) increased sensitivity to subtle phenotypes, which often can not be scored reliably by eye. The requisite automation of sample preparation and image acquisition has the welcome side effect of improving consistency and providing a permanent record of the experiment.

- **Creating open-source software for the *C. elegans* community**. *C. elegans* is used for studying complex multicellular biological processes by more than 11,000 researchers in 750 laboratories worldwide (http://www.WormBase.org, January 2010), and the close-knit community rapidly shares methods[17,18,34,35]. Based on our experience developing the CellProfiler software system (see Preliminary studies), packaging automated image analysis algorithms in user-friendly software encourages their use by the broader research community. Although we developed CellProfiler solely for high-throughput screening, 70% of studies citing it actually used it to quantify low-throughput assays (fewer than 100 samples). In this proposal we focus on developing algorithms that are robust and efficient for large-scale experiments, but we anticipate they will become an everyday tool for many researchers in the *C. elegans* community, a good investment since many of these are funded by the NIH.

Thus, in addition to the discovery of potential drugs and drug targets related to metabolism and infection, which could significantly impact the global burden of human disease, our aims will yield open-source software for automated, accurate, quantitative scoring for a wide range of *C. elegans* image-based assays that are currently intractable. The impact will be multiplied by *C. elegans* laboratories worldwide using the resulting software to study a wide variety of pathways relevant to basic biological research and human disease, in both low-throughput and high-throughput experiments.

B Innovation

In response to the strong demand for *C. elegans* screening, we propose to build on our technological innovations in sample preparation and imaging and our computational innovations for cells and brains to now create a novel technology for *C. elegans*. Our proposed work to develop novel algorithms for identifying and characterizing worms in microscopy images will bridge the final gap, for the first time enabling widespread identification of genetic and chemical regulators of human biological processes and diseases via whole-organism screening.

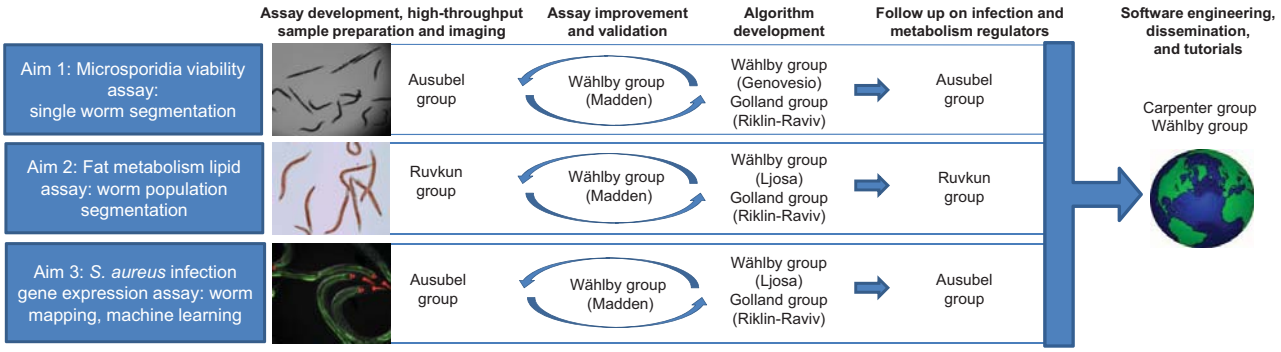


Figure C.1: Project overview, including the contributions of collaborating groups.

Automated image analysis for high-throughput screening of *C. elegans* is, in itself, novel: screens have so far been performed by eye due to the lack of suitable image analysis algorithms (excepting our simple *E. faecalis* screen[11]), limiting the number, types, and sizes of screens. Visual examination for a genome-wide RNAi screen takes 0.5–4 people-years; a large chemical screen[1] requires more than 10 people-years. Using the algorithms we will develop, such screens can be analyzed in weeks or months. Existing algorithms for *C. elegans* are insufficient; they were developed for low-throughput, high-resolution, 3-D, or time-lapse images[36–46], or for embryos, which have a different appearance than adults[47–53].

Several algorithmic innovations are necessary in order to quantify a variety of *C. elegans* phenotypes and attain the robustness required for routine high-throughput screening. We propose a novel, simplified representation for worm shapes that lends itself to a probabilistic interpretation. This allows us to adapt shape models to identification of worms in a high-throughput context, and leads to a novel algorithm for detangling worms by morphology-guided graph search. We will also build upon methods from our work in deformation analysis[54] and per-cell classification of cellular phenotypes by machine learning[55] to quantify phenotypic variation and fluorescence localization in individual worms.

C Approach

Overview of the team and the approach

The proposed project is founded on several multi-year existing collaborations between groups studying infection and metabolism using *C. elegans* (Ausubel and Ruvkun), and computational groups focused on developing algorithms for biomedical research (Wahlby, Carpenter, and Golland), making us uniquely situated to accomplish the proposed aims. As shown in Figure C.1, our interdisciplinary team is highly interactive and our approach to image assay development is a highly iterative process; typically the majority of the work is in multiple rounds of validation and testing of novel or existing algorithms while optimizing sample preparation protocols to ensure robust real-world performance. Each proposed aim is independent, but in several instances, improvements made for one aim will benefit the others. Later sections detail our proposed algorithm development for each aim, which will occur in the rich, collaborative, interdisciplinary environment of algorithm and software development at the Broad Institute and MIT. Here we outline the team and the approach.

Project leadership and algorithm development: The PI, Carolina Wahlby, will lead and coordinate the collaborating groups for the project. Based on Dr. Carpenter's work with Golland's group across the street at MIT's CSAIL (since 2004[55–59]) and the Ausubel and Ruvkun *C. elegans* laboratories across the Longfellow Bridge at MGH (since 2005[11,60]), Dr. Wahlby was able to quickly take leadership of these projects in 2009, start her own collaborations, and develop new ideas for *C. elegans* image analysis with the Golland group (see support letter). In less than one year, this collaboration resulted in a joint, peer-reviewed paper accepted for publication[61], another submitted, and the present proposal. The project's success so far is due to Dr. Wahlby's strong computational background and previous experience managing highly interdisciplinary collaborations on application-oriented image analysis (see Bengtsson and Ekström/Alderborn support letters).

Wet laboratory work: The Ausubel and Ruvkun groups are separately funded, equipped, and committed to completing the wet laboratory work to image thousands of samples for each assay (see Table C.1, Preliminary studies section, Resources file, and support letters). Furthermore, the laboratories are dedicated to the study of infection and metabolism and are separately funded to follow up on "hits" from the screens, in some cases

Table C.1: Overview of image sets to be collected by Ausubel and Ruvkun groups.

Aim	Assay	Group	Images	Scale
1	Viability assay (Microsporidia)[c]	Ausubel & Ruvkun	Brightfield only (shape reveals viability of worms in response to infection)	5000–100,000 chemicals
2	Lipid assay[b]	Ruvkun	Brightfield of oil red O (stains lipids)	Genome-wide RNAi
3	*S. aureus*-induced expression pattern[c]	Ausubel	Brightfield + GFP-fluorescence (reports expression of *clec-60:GFP* in response to infection) + *myo-2:mCherry*	Genome-wide RNAi

Screens funded by: [a]NIH R01 AI085581-01 [b]NIH R01 DK070147-06 & Broad Institute [c]NIH R01 AI064332-05 & R01 AI072508-02.

collaborating with the Broad Institute's Chemical Biology Platform, which has extensive experience in converting hit compounds into usable research probes or drugs.

Software development, dissemination, resource sharing, and reproducible research: The Carpenter group (see support letter) will implement, test, and disseminate the project team's algorithms into readily usable software following good software engineering practices. In keeping with the Broad Institute's mission to create advanced research tools for the scientific community, the Data/Software Sharing file details our plans for comprehensive sharing of both the data (images) and software produced. Specifically, the algorithms developed will be made readily usable by biologists via the open-source CellProfiler software project for high-throughput image analysis[57]. A major advantage of this system is that each analysis run retains complete information about the algorithms and settings used, enabling reproducible research[62]. CellProfiler runs on Windows, PC, and Unix systems, including computing clusters, and reads many image file formats via the BioFormats library[63]. The *C. elegans* algorithms will also be available via ImageJ[64], due to a funded project to interface it with CellProfiler (Carpenter, Eliceiri, and Rasband). Building on this existing software eliminates the waste of building a separate interface for worm algorithms and ensures longevity and dissemination for the algorithms.

In addition to software engineering for the project, the Carpenter group will also be primarily responsible for software dissemination and support through direct training with other high-throughput *C. elegans* laboratories (see Roy, Mylonakis, and Yanik support letters, for example), via conferences (e.g., The International *C. elegans* Meeting, Worm Genomics and Systems Biology Conference), via the Worm Breeder's Gazette[35], via online tutorials, and via public *C. elegans*-specific tutorials to train biologists to use the software.

Timeline: Work on Aim 1 will take place during the first two years. Work on Aim 2 will commence six months after funding and will be finished by the end of the third year. Work on Aim 3 will begin halfway through the second year and will be finished by the end of year 5.

Preliminary studies supporting the approach

In this section, we describe the independent and collaborative research completed within and among the Wahlby, Carpenter, Golland, Ausubel, and Ruvkun groups that provides the foundation for this proposal.

High-throughput *C. elegans* microscopy screen for regulators of *Enterococcus faecalis* infection: We recently published the first whole-animal *C. elegans* microscopy screen analyzed by automated image analysis[11]. Building on a smaller, manually-scored screen[65], we tested 37,214 chemicals for their ability to rescue *C. elegans* worms from an otherwise lethal *E. faecalis* infection. We acquired fluorescence images of the dead worms stained with SYTOX dye, plus brightfield images showing the entire worm population. Although the image-analysis approach was relatively simple, the screen uncovered six structural classes of compounds that are "anti-infectives" and appear to cure *C. elegans* animals without directly affecting the growth of *E. faecalis*. Three of these are novel structural classes of compounds that were not found in *in vitro* screens for antimicrobial compounds. This validates a major premise of our proposal, that image-based screens in the whole organism *C. elegans* will reveal compounds acting through novel mechanisms of action, in this case, mechanisms that are only manifest when the complex host/pathogen relationship is intact.

High-throughput *C. elegans* sample preparation, image acquisition, and assay development: The Ruvkun and Ausubel labs, with help from the Carpenter group, have established the pioneering *C. elegans* High-Throughput Screening Core Facility[66]. Both groups have extensive experience in developing assays and conducting large-scale screens to probe important biological questions in *C. elegans*, having completed manually-scored *C. elegans* screens relating to longevity[67–69], *E. faecalis* infection[65,70], metabolism[71,72], RNA inter-

ference[73], *Candida albicans* infection[74], synapses[75], immune response[76], molting[77], miRNA[78], diabetes[79], innate immune signaling[80], neuroendocrinology[81]. The specific assays they developed for this proposal are described later, in the context of each Aim.

The screening center uses a workflow in which a precise number of worms within a specified size/age range are dispensed by a COPAS large particle sorter into 4–6 multi-well plates per hour, and subsequently processed using automated plate washers and microscopes. The workflows enable both RNAi and chemical screening and imaging at multiple wavelengths. The team is skilled at optimizing assay parameters such as genetic background, readout, food source, salt concentration, temperature, timing, number of replicates, and number of animals per well. Imaging is optimized by transfer from agar to liquid media to minimize imaging artifacts and a paralytic drug is often added to slow worm movement, minimizing misalignment between subsequently imaged channels. Microscopy imaging is the primary screening method: plate readers do not offer per-worm or morphological readouts and are often not compatible even with bulk fluorescence-level assays[66]; customized flow cytometers can measure certain phenotypes[82,83], but current equipment to retrieve worms from a 96-well plate is too slow and inconsistent.

Both laboratories lead their fields and have productive records of pursuing hits from *C. elegans* screens. The Ausubel laboratory's reputation stems from pioneering discoveries that many human microbial pathogens also kill *C. elegans*[84–89], typically using similar virulence factors[84,86–92], and that key features are shared between *C. elegans*' immune system and the innate immune systems of mammals[80,93–97]. The Ruvkun lab is well-respected for work using *C. elegans* molecular genetics and genomics, leading to the discovery of microRNAs[98], the first detection of microRNAs in other animals[99], and the discovery of their role in gene regulation. Most relevant to the proposed project is the discovery of key members of the insulin pathway that control metabolism and longevity[100], that were later found to be conserved in mammals.

Development of image analysis and machine learning algorithms for biomedicine: The Wahlby, Carpenter, and Golland groups each have substantial experience developing and applying image analysis algorithms to important problems in cell biology and biomedical imaging. Our expertise spans the full spectrum required for the proposed project: developing advanced image analysis algorithms, validating them in the context of real-world biological problems, and creating practical, useful software tools that are made publicly and freely available.

Dr. Wahlby was one of the pioneers in developing advanced segmentation methods for phenotype quantification in fluorescence microscopy images of cells[101], using nuclear stains for seeded segmentation of cytoplasms[102], a widely used approach today. Our algorithms for accurate delineation of individual cells in culture and tissue[103] have proven valuable in a number of our own image-based biological experiments[104–108]. The algorithms have become widely used via a software tool[109] that also incorporates our novel algorithms for signal detection[110]. The algorithms are also a key component of CellProfiler[57] as a result of our collaboration with the Carpenter group in 2003. We have also developed a new apprach for quantification of signal colocalization[111,112] and designed methods for quantitative measurements using novel staining techniques[113,114]. Our recent work on *C. elegans* with this proposal's collaborators produced a novel method for segmentation of clusters of worms using a probabilistic shape model[61].

The Carpenter and Golland groups began collaborating in 2004 to expand the range of cell types and phenotypes amenable to automated analysis for high-throughput screening. This produced algorithms for the accurate identification of cell edges based on Voronoi diagrams in an image-based metric space[56], an approach for illumination correction for fluorescence microscopy images[58,115,116], a combination of existing algorithms, including Wahlby's, for the accurate identification of difficult-to-segment nuclei[57], a workflow for handling the unprecedented hundreds of numerical measurements for each of millions of cells in dozens of experiments[58], and a software infrastructure in which to incorporate these algorithms and approaches (detailed below). The software, CellProfiler, and some of its algorithms will be useful for the *C. elegans* work proposed here. Most importantly, these algorithms have been cited in hundreds of papers in the past three years, demonstrating that they serve an unmet need in biomedical research. We directly collaborated in many important studies in a wide variety of biological fields of study[55,57,59,117–126].

Machine learning has become increasingly useful in our work on scoring phenotypes in image-based screens in cases when a complex combination of features is required to differentiate between classes. We adapted the principles of content-based image retrieval[127] and created a system for scoring complex phenotypes in high-

throughput image-based screens using iterative feedback and machine learning[55]. We have used this software for large-scale screens for dozens of phenotypes that could not be scored by traditional methods[55,128], many of which are likely to be published in the next 1–2 years. Typically, no customization is required to accurately score phenotypes, aside from initial segmentation and feature extraction, overcoming a significant bottleneck in assay development for screens. The approach and software should be equally successful for *C. elegans* screens, once accurate measurements can be obtained from individual worms.

Aside from this collaborative work, the Golland group has established computational frameworks for image-based statistical analysis of shapes as well as shape-based segmentation. Their shape analysis research[129–133] explores the morphological variability of brain structures across and within different populations, which led to the development of a discriminative shape model[54]. The underlying mathematical frameworks are either level-set or MRF models—both are state-of-the-art techniques for segmentation. While the analysis and extraction of brain structures has been the main focus of the group's research[134], the segmentation of natural images with various forms of priors such as shape symmetry, GMM models and user interaction have also been explored[135–141]. These two complementary lines of research cover most aspects of the problems at hand—foreground/background segmentation, delineation of individual worms based on shape, and extraction of numerical measurements that are specific to worm phenotypes.

Modular open-source software for image analysis: Together, the Wahlby, Carpenter, and Golland groups have a track record of producing user-friendly software that is valued by the scientific community and capable of generating useful biological discoveries. The Carpenter, Golland, and Sabatini groups launched the open-source CellProfiler software project to give biologists a user-friendly interface to mix and match advanced image analysis algorithms (including our own, described above) in a modular way for high-throughput experiments[56–58,115,125,142–144]. We also created companion software, CellProfiler Analyst, for the exploration and analysis of multi-dimensional, image-based screening data which could not be handled by existing software, commercial or open-source[125]. These tools will be directly applicable to *C. elegans*-derived data.

CellProfiler has been useful to the biological community by many measures: (a) It has been cited more than 150 times in the 3 years since publication, including high-profile studies unaffiliated with our groups,[145–155] (b) The CellProfiler software is downloaded at a rate of 360/month, (c) There was widespread support from screening centers and laboratories around the world for our recent NIH R01 proposal to support CellProfiler.

Aim 1: Algorithms for *C. elegans* viability assays to identify modulators of pathogen infection

To score chemical perturbants for their ability to rescue *C. elegans* from an otherwise lethal infection by the pathogen Microsporidia, we will develop algorithms to count live and dead worms in each sample. These algorithms will delineate individual worms from clusters of worms and extract shape features that can distinguish curvy, live worms from straight, dead worms.

The successful *C. elegans* viability screen described in Preliminary studies[11] relied on measuring a fluorescent viability stain (SYTOX) across the population without needing to identify individual worms. However, for this Microsporidia assay, and other future live/dead screens, it is preferable to instead classify each animal as live or dead based on its shape in brightfield images; SYTOX staining adds reagent costs and sample preparation time and it is a less reliable indicator of viability from a biological perspective[66]. In addition, SYTOX stains some pathogens we plan to screen as well as some types of debris, thus obscuring the signal from the worms.

Experimental approach

While non-touching worms can usually be delineated in brightfield images based on the differences in intensities between foreground and background, image intensity alone is not sufficient for touching and overlapping worms. The high-throughput screening assays addressed here require algorithms that separate touching and overlapping worms in static images, where motion cues are unavailable. Moreover, edges and intensity variations within the worms often mislead conventional segmentation algorithms. On the other hand, while the varying postures of the worms introduce significant extrinsic geometrical differences, the worms have similar intrinsic geometrical properties (such as length and width profile). We propose a probabilistic shape model that captures this type of knowledge in an automated segmentation method. The key ideas are the construction of a low-dimensional shape-descriptor space and the definition of a probability measure on it. Closely related approaches for shape representation include the active shape model (ASM) and its variants[156], and medial

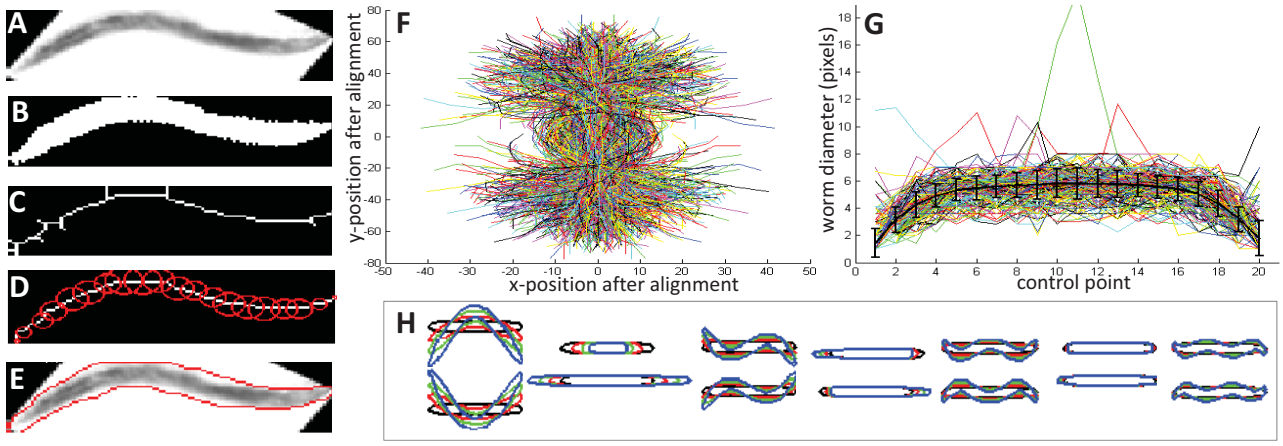


Figure C.2: Constructing a worm model. A: Rotated input image. B: Initial segmentation. C: Skeleton. D: Pruned skeleton with local radii at control points. E: Parameterized shape, recreated from descriptor. F: Connected control points of N=454 training worms after mirroring, alignment by translation, and rotation. G: Variation in radius along the length of all N worms. H: The effect of varying the weights of the "eigenworms" corresponding to the seven largest eigenvalues of the final model. Any 3-D properties of the worms will be captured as projections in our 2-D images. In fact, the sixth "eigenworm" appears to capture the C. elegans *lifting its head.*

axis transform methods[157] for capturing shape variability in anatomical structures and other objects[158,159] and others. We learn the possible shape variations from N training worms obtained by automated segmentation of a subset of worms that do not touch or overlap.

1. Construct a low-dimensional worm shape descriptor from the skeleton of the shape and its distances to the boundaries, given by the medial-axis transform[160]. Fig. C.2 exemplifies our proposed computationally efficient representation of the shape, where we extract the skeleton of each worm (Fig. C.2C), and prune spurs by iteratively removing the shortest spur of every branch point of the skeleton. Once a non-branched skeleton is obtained, we find end points, and sample n control points uniformly along the skeleton. The original worm shape can be approximately restored by placing discs with radius equal to the local worm width (Fig. C.2D) at each control point, and smoothing the edges by the pair-wise convex hull of the discs (Fig. C.2E).

2. Reduce dimensionality by Principal Component Analysis (PCA): Align descriptors by similarity transformation (i.e., rotation and translation, no scale or skew) by minimizing the sum of the Euclidean distances of corresponding points along the skeletons. Thus, non-rigid components of the deformations are completely captured within the shape variations. To make variations in worm shape symmetrical, the training set is doubled to $2N$ by mirroring all samples. Fig. C.2F shows the aligned skeletons of the training set. The significant similarity of the worms' radii profiles (Fig. C.2G) allows representation of the differences in the radii by a single value, which corresponds to the median thickness of the worm. The deformations of the postures are described by the coordinates of the n aligned control points and the variation in thickness, resulting in a $(2n+1)$ dimensional data space. We project the vector representations of the parameterized skeletons into a lower-dimensional feature space by PCA[161]. All the worms in the training set can be restored with good approximation by linear combinations of the eigenvectors, or "eigenworms" (Fig. C.2H).

3. Find posture probabilities and resolve clusters by graph search algorithm: The weights $\mathbf{w}$ of the training worms define a probability measure on the feature space of the worm deformations: $p(\mathbf{x}) \propto \exp(-\mathbf{w}^T \Sigma_L^{-1} \mathbf{w})$, where $\Sigma_L = \mathrm{diag}(\lambda_1 \ldots \lambda_L)$ as in[162]. After the input images have been partitioned into worm regions (individuals and clusters) and background (Fig. C.3B) as discussed later, we find the skeleton of each clusters using the medial-axis transform[160] (Fig. C.3C). We represent the skeleton by a sparse direct graph $G_s = \{V, E\}$. The vertices V of the graph represent the skeleton segments (Fig. C.3C)) and the edges E connect pairs of vertices representing pairs of skeleton segments with common intersection points. We represent a worm candidate by a path $p_1 \ldots p_N$ in the graph containing one or more vertices. Set K to the estimated number of worms in a cluster (given by cluster area) and Let $p_1 \ldots p_N$ denote the paths in the graph. We find K out of N paths in the graph by minimizing the cost functional

$$E(\mathbf{p}_1 \ldots \mathbf{p}_K) = -\sum_{k=1}^{K} \log P(\mathbf{p}_k) + \alpha \sum_{k=2}^{K} \sum_{l=1}^{k-1} |\mathbf{p}_k \cap \mathbf{p}_l| + \beta |\tilde{V}_k|,$$

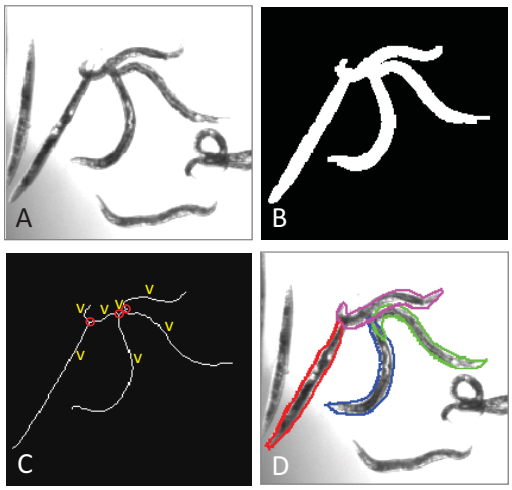


*Figure C.3: Resolving clusters. A: Input image, B: binary image of cluster, and C: its pruned skeleton: **v**s indicate vertices; **o**s indicate groups of edges. D: Final segmentation result.*

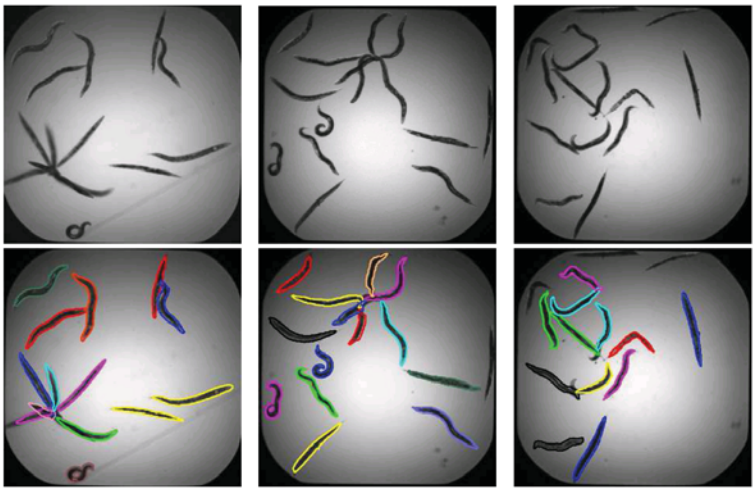


Figure C.4: Three examples of resolved clusters (bottom) shown together with original images (top). Worms close to the well edge were excluded from this analysis.

where $|\cdot|$ denote cardinality or size. The first term is a requirement that the selected set of paths will have the highest probability to represent true worm shapes. The second term is a requirement that the sum of pairwise overlaps between the selected paths will be minimal. The third term is the number of vertices that are not included in the union of the selected paths, constraining the paths to cover the worm-cluster skeleton, and α and β are scaling factors. A global minimum can by obtained by an exhaustive search for all the subsets of K out of N paths in G_s. This is however a combinatorial problem of order $\binom{N}{K}$. To reduce the computational time we apply a greedy[163] strategy where at each stage we make a locally optimized choice of a path in the graph, until we select K paths. We applied the proposed segmentation approach to images containing worm clusters that could not be resolved based on gray-scale information alone. Most of the worms were correctly segmented as verified by visual evaluation (Fig. C.3D and C.4).

4. Measure worm viability by scoring the live/dead phenotype as the worm's length along the medial axis divided by the straight distance between worm's end points[37,164]. Initial studies also indicate that the shape characteristics described by the eigenworms provide a good measure of viability.

Validation, evaluation, and benchmarks

To validate and evaluate the proposed algorithm we will use a set of 6000 expert-annotated brightfield images from a previous screen[11] in addition to images from the Microsporidia screen itself. Overall, our goal is to achieve "screenability" in terms of both accuracy and computational speed. ***Accuracy***: We will use metrics accepted in the screening field to assess accuracy based on the ability to distinguish control wells with worm populations of known phenotype—hundreds of these controls are included in each experiment. If the assay readout is Gaussian, we will aim for a Z'-factor[165] above 0.5 ($>$0.2 would still be acceptable); if not, we will use classification sensitivity and specificity, overall aiming to avoid visual examination for 90–95% of the samples. During the iterative process of algorithm and assay development, we will also validate individual steps of the image analysis pipeline (foreground/background segmentation, worm cluster resolution, live/dead scoring) as appropriate, comparing algorithm results to "ground truth" provided by our worm experts. ***Speed***: Image processing should keep pace with image acquisition; given current image acquisition rates and cluster computing costs, our goal is 6 CPU-minutes or less per image on a typical CPU. The methods proposed are likely to meet this goal, but there are many ways to reduce computational costs if needed.

Potential problems and alternative strategies

Initial foreground/background segmentation is a prerequisite for the proposed cluster separation. If local adaptive thresholding is not sufficient, we will rely on more advanced methods, such as level-sets for foreground/background separation (Aim 2).

Cluster skeletonization may not coincide with the centers of the worms, skewing the cluster separation.

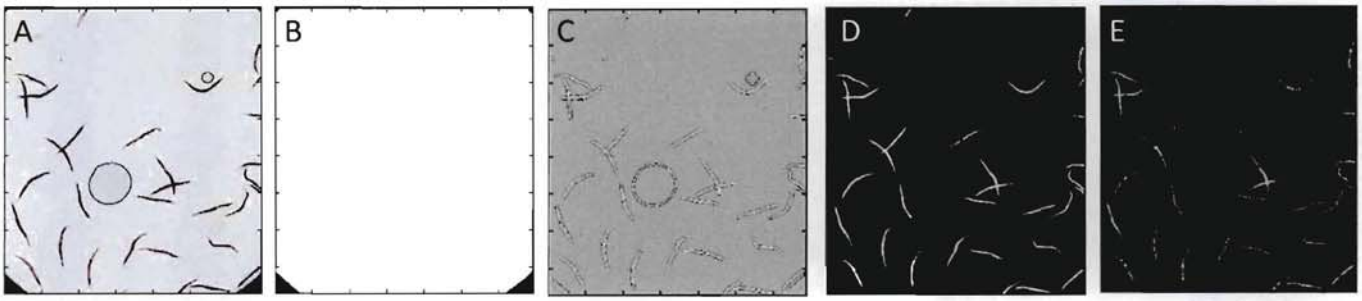


Figure C.5: A: Input image with well edges and bubbles of size and color similar to the worms. B: The well edges masked away by convex hull. C: Gradient magnitude defines bubble edges. D: Inverted image after removal of artifacts. E: Outline of worms (green) after background illumination correction.

Over-segmenting the clusters using watersheds[166] will, apart from dividing the worms into many pieces, also place watersheds at bright ridges between worms. Merging will not entangle crossing worms, but a selective merging that keeps watershed boundaries placed at bright ridges (based on local intensity information, similar to our previous work[103], but allowing also incomplete watersheds), will lead to a binary image where bright ridges are marked as background. A skeleton of such an image is more likely to guide the probabilistic shape model to a correct segmentation result. A distance transform of the binary image can guide the merging step, forcing it to preserve ridges located at a worm's thickness from the cluster edge.

Scoring viability from clusters: If individual worms cannot be segmented, we will measure the proportion of cluster area occupied by straight worm segments by a simple algorithm that fits long line segments inside the cluster. The algorithm considers all pairs of pixels in a connected region, and if a line>75% of the typical worm length can connect the pair while remaining in the worm region, the pixels along the line are marked as belonging to a dead worm.

Aim 2: Algorithms for *C. elegans* lipid assays to identify genes that regulate fat metabolism

To identify regulators of fat metabolism, we will extract lipid-related phenotypic features from populations of worms. This requires robust foreground/background separation, artifact removal, and definition of biologically relevant feature descriptors. Improvements in the first two of these goals will be applicable to a variety of assays, including those described in Aims 1 and 3.

The Ruvkun group completed a genome-wide *C. elegans* RNAi screen for genes regulating lysosomal content using the fluorescent dye Nile Red[60,71], revealing a wide range of functional components of the mammalian cellular wasting cycle, due to the conservation between *C. elegans* and mammals in these pathways. Although it was the first screen to probe these pathways in an intact, living animal, the scoring was manual and non-quantitative. The group recently discovered that the stain oil red O, unlike Nile Red, labels the major fat storage compartment[60]. We expect to uncover novel regulators of energy metabolism using this true fat stain. We will carry out the screen using a *C. elegans* strain hypersensitive to RNAi. We will also perform the screen in *C. elegans* strains with perturbations in metabolic/longevity pathways; in the insulin-signaling deficient mutant *daf-2* and the calorically restricted mutant *eat-2*. We have already acquired images from >4,000 samples (in duplicate) after many months of iterative improvements in sample preparation and image acquisition.

Experimental approach

As compared to Aim 1's viability assay, which requires identification of individual worms to measure shape, the lipid assay can be scored by averaged measurements from non-separated worms. The challenges include robust separation of image foreground (worms) from background, elimination of well edges and artifacts, and identification of descriptive features that reflect the fatness phenotype of each worm population correctly. We have discovered that identifying the foreground in brightfield images of *C. elegans* requires a more accurate intensity threshold than is the case for most fluorescently labeled cell-based assays. In some cases, local adaptive thresholding[167] is sufficient, but for more difficult cases we propose to define foreground/background using level-set segmentation that combines image intensity with gradient information. The accidental inclusion of non-worm material in the segmentation result may skew extracted feature measures leading to poor accuracy.

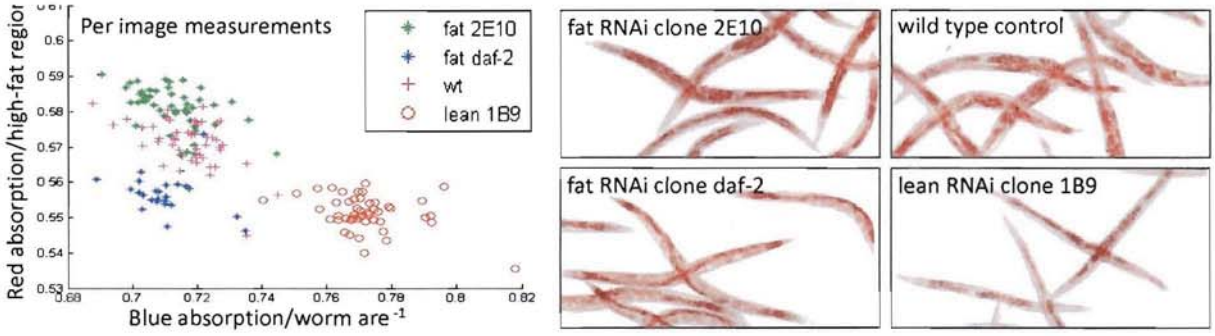


Figure C.6: Quantification of oil red O accumulation (an indicator of fat accumulation) for worms fed with different bacterial RNAi clones. Each point in the plot represents a worm population in a well.

Due to the specifics of sample preparation, this assay produces many artifacts that are not seen in Aim 1's assay, such as bubbles. We propose to eliminate artifacts based on gradient magnitude and color.

1. Correct for background illumination variations using existing methods based on iteratively fitting a surface (B-spline) to the image background[168]. Illumination correction of color images is often performed on the L-component in L*a*b color space;[167] here we propose to apply the correction on each of the RGB-color channels separately, giving the combined effect of illumination and color correction in a single step.

2. Segment foreground/background by level-set methods that rely on low level image data (intensity and gradients). We use the level-set formulation[169] for a parameterization-free representation of the evolving worms contours. We define segmentation by assigning the positive and the negative levels of the level-set function to the object foreground and background respectively, representing the object boundary by the intersection of the level-set function with the zero plane. The functional consists of a region-based term that encourages homogeneity of image intensity in semantically related regions[170,171], an edge-based term that rewards coincidence of the object boundary with the image edges[172,173], and an edge alignment constraint that encourages evolution of the object boundary in a direction normal to the image edges[174,175]. We use the first variation of the level-set functional to define the gradient-descent equations that control the evolution of the object contour.

3. Identify and eliminate image artifacts starting with the well edge, which can often be found by simple intensity thresholding. However, using a binary thresholding result as a mask often leads to the loss of dark objects close to the well edge. We therefore refine the thresholding result by defining the convex hull[167] of the binary mask, leading to a simple and robust well segmentation (Fig. C.5B). We filter out bubbles based on their gradient magnitude being greater than that of the worm edges (see Fig. C.5A vs D). Finally, we filter out other artifacts such as bacterial aggregates and dust based on color, texture, and median of a distance transformation[167], which provides a valuable metric to discriminate between worm clusters and artifacts based on thickness.

4. Extract fat-related features from populations of worms such as averages of intensity, texture and color (defined by individual color channels and their ratios). We will also extract normalized granularity features[176] to quantify the texture of the oil red O stain, and measure worm width without extracting individual worms by distance transformation of the binary image foreground. The distribution of values within the resulting distance map provides a measure of width, discriminating between populations of thick versus thin worms.

5. Select descriptive features by machine learning applied to a large number of features extracted from set of control images with known phenotypes, prepared in parallel with the screen images (see Aim 3, although there in the context of per-worm measurements). Fig. C.6 shows four worm phenotypes from a preliminary experiment of the lipid assay. The two features that resulted in the best separation of the four phenotypes were extracted and plotted (Fig. C.6), where each point represents an image of a population of worms in a well, yielding preliminary separation of some phenotypes.

Validation, evaluation, and benchmarks: Our approach to evaluate screenability is the same as in Aim 1.

Potential problems and alternative strategies

Difficult foreground/background segmentation: [redacted]

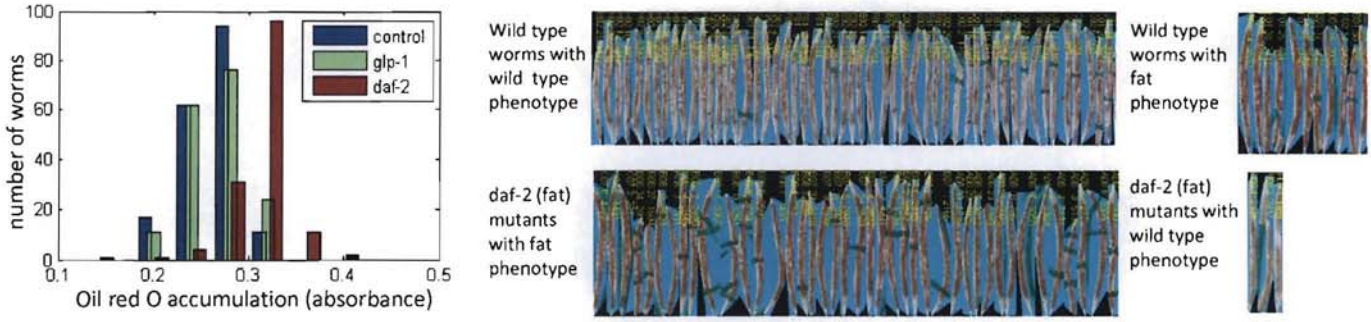


Figure C.7: Worm populations are often heterogeneous making measurements extracted from individual worms more powerful to identify phenotypes.

[redacted]

Worm population heterogeneity from non-penetrant RNAi may lead to poor separability of phenotypes, as they disappear in population averages. If this is the case, feature extraction from individual worms will likely be more powerful. Fig. C.7 shows data from an initial experiment where oil red O stain was quantified in individual worms. Although genetically identical, some wild type worms display the phenotype of fat mutants, and vice versa. The assay may be best scored by the percentage of individual worms meeting threshold criteria. Per worm measurements can be extracted by selecting only those worms in the population that do not touch or overlap (as long as this does not introduce a bias), or by using the methods proposed in Aim 1. We will also investigate a prior-based level-set approach[135,136,139,177–181] for the extraction of individual worms incorporating the shape model described in Aim 1.

Subtle phenotypes: Anatomical information, such as localization of fat to worm embryos or gut, plays an important role in visual phenotype interpretation. If the features described above are not sufficient to discriminate between subtle phenotypes, we will apply the anatomical atlas described in Aim 3.

Aim 3: Algorithms for gene expression pattern assays to identify regulators of the response of the *C. elegans* host to *Staphylococcus aureus* infection

To identify regulators of an animal's response to infection by the clinically important human pathogen *S. aureus*, and to identify potential anti-infectives with novel mechanisms of action, we will use machine learning to classify relevant infection-response phenotypes based on the staining patterns in individual worms.

The Ausubel group has fused the promoter of *clec-60* to GFP to create a transgenic strain of *C. elegans* that expresses the GFP reporter only when infected with *S. aureus* (Fig. C.8A). Expression is normally constrained to the posterior intestinal cells upon infection; our goal is to identify samples where the immune-related pathways have been perturbed and the pattern is altered; the targets of these perturbations will be regulators of the expression of immune effectors.

Experimental approach

We will measure signal localization and local texture after subdividing the animal in two different ways, then use machine learning to discern the phenotype of interest. The proposed steps are as follows:

1. Map worms to a canonical coordinate system where variations in posture have been removed while minimizing the deformation of the textures and intensity distributions within the worms. This is a prerequisite for comparing localization patterns between worms that are posed differently. It is also valuable for visual examination, as a montage of straightened worms provides a clear visual overview that can help validate hits.

As depicted in Fig. C.8, we will extract each single worm (or worm in a cluster, using the segmentation techniques of Aim 1), then re-map each worm by extracting a series of one-pixel–separated lines orthogonal to the medial axis curve and align them along a straight line that represents the anterior-to-posterior extension of the worm (Fig. C.8D–F). Distinguishing head from tail is facilitated by the use of a *C. elegans* strain whose head is labeled with the red fluorescent marker *mye-2:mCherry*.

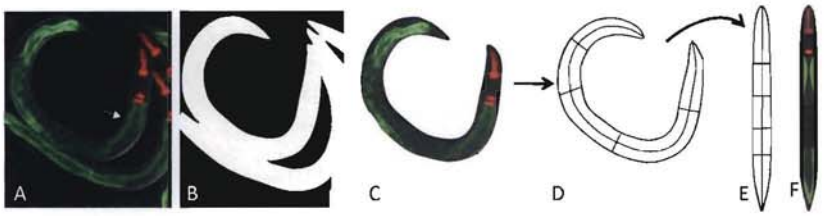


Figure C.8: Worm straightening (schematic). A: Original fluorescent image. B: Foreground/background segmentation of corresponding brightfield image. C: Single worm from fluorescence image. D: Medial axis and cross sections, straightened in E and re-sampled to F.

Previous work on worm straightening has established that the main problem is that of finding a medial axis curve in 2D and 3D[36]. For the limited resolution of HTS images, we propose to use control points along the medial axis transform[157] (as in Aim 1), and approximate the worm's medial axis by a smooth spline function. We expect the method to successfully minimize the loss of intra-worm morphology because the rotations of the lines will be rigid, and the only loss of image resolution will be due to pixel interpolation at rotation.

2. Extract intensity and texture features from fixed regions along the worm to capture the spatial aspects of the localization. We will partition the straightened worm into pieces of equal length, then extract intensity and texture features from each partition.

3. Create an atlas of the main anatomical features to allow features to be extracted from biologically meaningful regions. In this assay, the emphasis is on signal localization and texture; a worm is not a random structure, and by mapping each worm to an anatomical atlas simplified to the resolution available in a high-throughput experiment, the context of the signals can be accessed.

We will compute the mean length and the mean width profile from a representative subset of the straightened worms in the experiment. The Ausubel lab will inspect 5–10 randomly sampled worms and independently outline the epithelium, gut, head, and uterus. These are the anatomical features that are easily discernible with the magnification and staining of this assay; for other assays the regions of interest will be different. Previous work has demonstrated the feasibility of constructing a worm atlas for high-resolution 3-D images[182]. Some of its measurement techniques may transfer, but different algorithms are required for low-resolution, high-throughput screens where multiple worms touch and overlap.

4. Extract intensity and texture features from each region in the atlas. We will deform the atlas to match the shape of the straightened worm, then extract intensity and texture features from the regions corresponding to each anatomical feature. We plan to deform the atlas rather than the worm because it will be beneficial to preserve the scale of the textures in the tissue.

5. Train a boosting classifier to discern the phenotype of interest in individual worms. We will train the classifier in an iterative fashion, building on techniques that have performed well for cells[55]. Using an interactive software tool, our collaborators in the Ausubel lab will initially identify a few animals as positive or negative for the phenotype. The tool will then train a classifier and display a number of worms with putative labels. Next, our collaborators will correct the computer's errors, the classifier will be retrained, and so on. This iterative process continues until the classifier is sufficiently accurate.

We will use fast gentle boosting[183], which has performed well and shown resilience to overtraining in cell-based screens[55,128]. The resulting classifier consists of a sequence of rules (decision stumps), each of which is a nonlinear function of only one feature. Thus, it is more transparent than many other methods: the user can see which features the classifier is using. Once we have obtained boosting scores for each individual worm, we will compute enrichment scores for each sample by reference to a beta-binomial model fitted to the experiment-wide distribution of per-worm scores[55]. Based on our experience with subtle cellular phenotypes, we believe that a training set of a few hundred worms will be sufficient.

Validation, evaluation, and benchmarks: Our approach to evaluate screenability is the same as in Aim 1.

Potential problems and alternative strategies

Discriminating head from tail based on their width profiles as has been successful in previous work on high-resolution images of individual worms[164]. We propose this solution if no flurescent markers are available.

Overlapping worms: Worms are transparent, so the signals from two overlapping worms will mix. If this becomes a problem, we will algorithmically mark overlapping pixels before straightening and transformation so that they can be excluded from the feature extraction steps.

Bibliography & References Cited

[1] Austin CP, Brady LS, Insel TR, Collins FS (2004) Molecular biology: NIH molecular libraries initiative. Science 306: 1138–1139, URL http://dx.doi.org/10.1126/science.1105511.

[2] Collins FS (2010) Opportunities for research and NIH. Science 327: 36–37, URL http://dx.doi.org/10.1126/science.1185055.

[3] Eggert US, Mitchison TJ (2006) Small molecule screening by imaging. Curr Opin Chem Biol 10: 232–237, URL http://dx.doi.org/10.1016/j.cbpa.2006.04.010.

[4] Carpenter AE (2007) Image-based chemical screening. Nat Chem Biol 3: 461–465, URL http://dx.doi.org/10.1038/nchembio.2007.15.

[5] Ljosa V, Carpenter AE (2009) Introduction to the quantitative analysis of two-dimensional fluorescence microscopy images for cell-based screening. PLoS Comput Biol 5: e1000603+, URL http://dx.doi.org/10.1371/journal.pcbi.1000603.

[6] Inglese J, Johnson RL, Simeonov A, Xia M, Zheng W, Austin CP, Auld DS (2007) High-throughput screening assays for the identification of chemical probes. Nature Chemical Biology 3: 466–479, URL http://dx.doi.org/10.1038/nchembio.2007.17.

[7] Perrimon N, Mathey-Prevot B (2007) Applications of high-throughput RNA interference screens to problems in cell and developmental biology. Genetics 175: 7–16, URL http://dx.doi.org/10.1534/genetics.106.069963.

[8] Moffat J, Sabatini DM (2006) Building mammalian signalling pathways with RNAi screens. Nature Reviews Molecular Cell Biology 7: 177–187, URL http://dx.doi.org/10.1038/nrm1860.

[9] Boutros M, Ahringer J (2008) The art and design of genetic screens: RNA interference. Nat Rev Genet 9: 554–66, URL http://dx.doi.org/10.1038/nrg2364.

[10] Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL (2006) Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nature Reviews Drug Discovery 6: 29–40, URL http://dx.doi.org/10.1038/nrd2201.

[11] Moy TI, Conery AL, Larkins-Ford J, Wu G, Mazitschek R, Casadei G, Lewis K, Carpenter AE, Ausubel FM (2009) High-throughput screen for novel antimicrobials using a whole animal infection model. ACS Chemical Biology 4: 527–533, PMCID PMC2745594, URL http://dx.doi.org/10.1021/cb900084v.

[12] Zon LI, Peterson RT (2005) In vivo drug discovery in the zebrafish. Nat Rev Drug Discov 4: 1474–1776, URL http://dx.doi.org/10.1038/nrd1606.

[13] Chi K (2009) How to reel in high-throughput results using worms and fish. The Scientist 23: 51.

[14] Artal-Sanz M, de Jong L, Tavernarakis N (2006) *Caenorhabditis elegans*: A versatile platform for drug discovery. Biotechnol J 1: 1405–1418, URL http://dx.doi.org/10.1002/biot.200600176.

[15] Kamath RS, Ahringer J (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods 30: 313–321.

[16] Labbe JC, Roy R (2006) New developmental insights from high-throughput biological analysis in *Caenorhabditis elegans*. Clin Genet 69: 306–314, URL http://view.ncbi.nlm.nih.gov/pubmed/16630163.

[17] Wormbook (2008) The *C. elegans* research community. http://www.wormbook.org, URL http://dx.doi.org/10.1895/wormbook.1.7.1.

[18] Harris TW, Chen N, Cunningham F, Tello-Ruiz M, Antoshechkin I, Bastiani C, Bieri T, Blasiar D, Bradnam K, Chan J, Chen CK, Chen WJ, Davis P, Kenny E, Kishore R, Lawson D, Lee R, Muller HM, Nakamura C, Ozersky P, Petcherski A, Rogers A, Sabo A, Schwarz EM, Van Auken K, Wang Q, Durbin R, Spieth J, Sternberg PW, Stein LD (2004) WormBase: a multi-species resource for nematode biology and genomics. Nucl Acids Res 32: D411–417, URL http://dx.doi.org/10.1093/nar/gkh066.

[19] Sulston JE, Horvitz HR (1977) Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. Developmental Biology 56: 110–156, URL http://dx.doi.org/10.1016/0012-1606(77)90158-0.

[20] Sulston JE, Schierenberg E, White JG, Thomson JN (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Developmental Biology 100: 64 – 119, URL http://dx.doi.org/10.1016/0012-1606(83)90201-4.

[21] Burns AR, Kwok TC, Howard A, Houston E, Johanson K, Chan A, Cutler SR, Mccourt P, Roy PJ (2006) High-throughput screening of small molecules for bioactivity and target identification in *Caenorhabditis elegans*. Nat Protoc 1: 1906–1914, URL http://view.ncbi.nlm.nih.gov/pubmed/17487175.

[22] Hertweck M, Baumeister R (2005) Automated assays to study longevity in *C. elegans*. Mechanisms of Ageing and Development 126: 139–145, URL http://dx.doi.org/10.1016/j.mad.2004.09.010.

[23] Gunsalus K (2008) A *Caenorhabditis elegans* genetic-interaction map wiggles into view. Journal of Biology 7: 8, URL http://dx.doi.org/10.1186/jbiol70.

[24] Didier ES, Weiss LM (2006) Microsporidiosis: current status. Curr Opinion Infect Dis 19: 485–492.

[25] Williams BA (2009) Unique physiology of host-parasite interactions in microsporidia infections. Cellular Microbiology 11: 1551–1560, URL http://dx.doi.org/10.1111/j.1462-5822.2009.01362.x.

[26] Troemel ER, Flix MA, Whiteman NK, Barrire A, Ausubel FM (2008) Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. PLoS Biol 6: e309, URL http://dx.doi.org/10.1371/journal.pbio.0060309.

[27] Kurz C, Ewbank J (2003) *Caenorhabditis elegans*: An emerging genetic model for the study of innate immunity. Nat Rev Genet 4: 380–390.

[28] Huang ES, Basu A, O'Grady M, Capretta JC (2009) Projecting the future diabetes population size and related costs for the U.S. Diabetes Care 32: 2225–2229, URL http://dx.doi.org/10.2337/dc09-0459.

[29] Elston DM (2007) Community-acquired methicillin-resistant *Staphylococcus aureus*. Journal of the American Academy of Dermatology 56: 1–16, URL http://dx.doi.org/10.1016/j.jaad.2006.04.018.

[30] Garcia-Lara J, Needham AJ, Foster SJ (2005) Invertebrates as animal models for *Staphylococcus aureus* pathogenesis: a window into host-pathogen interaction. FEMS Immunology & Medical Microbiology 43: 311–323, URL http://www.wormbase.org/db/misc/paper?name=WBPaper00025121.

[31] Tarca AL, Carey VJ, Chen Xw, Romero R, Drghici S (2007) Machine learning and its applications to biology. PLoS Comput Biol 3: e116, URL http://dx.doi.org/10.1371/journal.pcbi.0030116.

[32] Zhang S, Jin G, Zhang XSS, Chen L (2007) Discovering functions and revealing mechanisms at molecular level from biological networks. Proteomics 7: 2856–2869, URL http://dx.doi.org/10.1002/pmic.200700095.

[33] Loging W, Harland L, Williams-Jones B (2007) High-throughput electronic biology: mining information for drug discovery. Nat Rev Drug Discov 6, URL http://dx.doi.org/10.1038/nrd2345.

[34] Antoshechkin I, Sternberg PW (2007) The versatile worm: Genetic and genomic resources for *Caenorhabditis elegans* research. Nature Reviews Genetics 8: 518–532, URL http://dx.doi.org/10.1038/nrg2105.

[35] Worm Book Consortium (2010) The worm breeders gazette. URL http://www.wormbook.org/wbg.html.

[36] Peng H, Long F, Liu X, Kim SK, Myers EW (2008) Straightening *Caenorhabditis Elegans* images. Bioinformatics 24: 234–242, URL http://dx.doi.org/10.1093/bioinformatics/btm569.

[37] Feng Z, Cronin CJ, John H Wittig J, Sternberg PW, Schafer WR (2004) An imaging system for standardized quantitative analysis of *C. elegans* behavior. BMC Bioinformatics 5, URL http://dx.doi.org/10.1186/1471-2105-5-115.

[38] Cronin CJ, Feng Z, Schafer WR (2006) Automated imaging of *C. elegans* behavior. Methods in Molecular Biology, *C elegans* 351: 241–251, URL http://dx.doi.org/10.1385/1-59745-151-7:241.

[39] Hoshi K, Shingai R (2006) Computer-driven automatic identification of locomotion states in *Caenorhabditis elegans*. Journal of Neuroscience Methods 115: 355–363, URL http://dx.doi.org/10.1016/j.jneumeth.2006.05.002.

[40] Cronin CJ, Mendel JE, Mukhtar S, Kim YM, Stirbl RC, Bruck J, Sternberg PW (2005) An automated system for measuring parameters of nematode sinusoidal movement. BMC Genetics 6, URL http://dx.doi.org/10.1186/1471-2156-6-5.

[41] Fontaine E, Barr AH, Burdick JW (2007) Model-based tracking of multiple worms and fish for biological studies. In: Proceedings of the Workshop on Dynamical Vision at the Eleventh IEEE International Conference on Computer Vision.

[42] Fontaine E, Burdick J, Barr A (2006) Automated tracking of multiple *C. elegans*. In: Proceedings of the 28th IEEE EMBS Annual International Conference, pp. 3716–3719.

[43] Tsechpenakis G, Bianchi L, Metaxas D, Driscoll M (2008) A novel computational approach for simultaneous tracking and feature extraction of *C. elegans* populations in fluid environments. IEEE Transactions on Biomedical Engineering 55: 1539–1549, URL http://dx.doi.org/10.1109/TBME.2008.918582.

[44] Ramot D, Johnson BE, Tommie L Berry J, Carnell L, Goodman MB (2008) The parallel worm tracker: A platform for measuring average speed and drug-induced paralysis in nematodes. PLoS One 3: e2208, URL http://dx.doi.org/10.1371/journal.pone.0002208.

[45] de Bono M, Bargmann C (1998) Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 94: 679–689.

[46] Zhou B, Baek J (2009) Toward automatic behavioral screen: A computational model for analyzing *Caenorhabditis elegans* locomotion. In: IEEE Eng Med Biol Soc, pp. 4937–40.

[47] Zipperlen P, Fraser AG, Kamath RS, Martinez-Campos M, Ahringer J (2001) Roles for 147 embryonic lethal genes on *C. elegans* chromosome I identified by RNA interference and video microscopy. EMBO Journal 20: 3984–3992, URL http://view.ncbi.nlm.nih.gov/pubmed/11483502.

[48] Gönczy P, Echeverri C, Oegema K, Coulson A, Jones SJ, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E, Hannak E, Kirkham M, Pichler S, Flohrs K, Goessen A, Leidel S, Alleaume AM, Martin C, Ozlü N, Bork P, Hyman AA (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature 408: 331–336, URL http://dx.doi.org/10.1038/35042526.

[49] Sönnichsen B, Koski L, Walsh A, Marschall P, Neumann B, Brehm M, Alleaume AM, Artelt J, Bettencourt P, Cassin E, Hewitson M, Holz C, Khan M, Lazik S, Martin C, Nitzsche B, Ruer M, Stamford J, Winzi M, Heinkel R, Röder M, Finell J, Häntsch H, Jones SJM, Jones M, Piano F, Gunsalus KC, Oegema K, Gönczy P, Coulson A, Hyman AA, Echeverri CJ (2005) Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 434: 462–469, URL http://dx.doi.org/10.1038/nature03353.

[50] Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH (2006) Automated cell lineage tracing in *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences of the United States of America 103: 2707–2712, URL http://dx.doi.org/10.1073/pnas.0511111103.

[51] Piano F, Schetter AJ, Morton DG, Gunsalus KC, Reinke V, Kim SK, Kemphues KJ (2002) Gene clustering based on rnai phenotypes of ovary-enriched genes in *C. elegans.* Current Biology 12: 1959–1964, URL http://dx.doi.org/10.1016/S0960-9822(02)01301-5.

[52] Bot NL, Tsai MC, Andrews RK, Ahringer J (2003) TAC-1, a regulator of microtubule length in the *C. elegans* embryo. Current Biology 13: 1499–1505, URL http://dx.doi.org/10.1016/S0960-9822(03)00577-3.

[53] Dzyubachyk O, Jelier R, Lehner B, Niessen W, Meijering E (2009) Model-based approach for tracking embryogenesis in *Caenorhabditis elegans* fluorescence microscopy data. In: IEEE Eng Med Biol Soc, pp. 5356–9.

[54] Golland P, Grimson W, Shenton M, Kikinis R (2001) Deformation analysis for shape based classification. In: IPMI, pp. 517–530.

[55] Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver SJ, Grenier JK, Castoreno AB, Eggert US, Root DE, Golland P, Sabatini DM (2009) Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc Nat Acad Sci USA 106: 1826–1831, PMCID PMC2634799, URL http://dx.doi.org/10.1073/pnas.0808843106.

[56] Jones TR, Carpenter A, Golland P (2005) Voronoi-based segmentation of cells on image manifolds. In: Proceedings of the First International Workshop on Computer Vision for Biomedical Image Applications (CVBIA), pp. 535–543, URL http://dx.doi.org/10.1007/11569541_54.

[57] Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, Golland P, Sabatini DM (2006) CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. Genome Biology 7: R100, PMCID PMC1794559, URL http://dx.doi.org/10.1186/gb-2006-7-10-r100.

[58] Jones T, Carpenter A, Sabatini D, Golland P (2006) Methods for high-content, high-throughput image-based cell screening. In: Proceedings of the Workshop on Microscopic Image Analysis with Applications in Biology (MIAAB), pp. 65–72.

[59] Jones TR, Kang IH, Wheeler DB, Lindquist RA, Papallo A, Sabatini DM, Golland P, Carpenter AE (2008) CellProfiler Analyst: Data exploration and analysis software for complex image-based screens. BMC Bioinformatics 9: 482, PMCID PMC261443, URL http://dx.doi.org/10.1186/1471-2105-9-482.

[60] O'Rourke E, Soukas A, Carr C, Ruvkun G (2009) *C. elegans* major fats are stored in vesicles distinct from lysosome-related organelles. Cell Metabolism 10: 430–435.

[61] Wählby C, Riklin-Raviv T, Ljosa V, Conery AL, Golland P, Ausubel FM, Carpenter AE (2010) Resolving clustered worms via probabilistic shape models. In: Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI), Rotterdam, The Netherlands, accepted for publication.

[62] Kovačević J (2007) How to encourage and publish reproducible research. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. IV–1273–IV–1276, URL http://dx.doi.org/10.1109/ICASSP.2007.367309.

[63] Swedlow JR, Goldberg IG, Eliceiri KW, The OME Consortium (2009) Bioimage informatics for experimental biology. Annual Review of Biophysics 38: 327–346, URL http://dx.doi.org/10.1146/annurev.biophys.050708.133641.

[64] Abràmoff MD RS Magalhaes PJ (2004) Image processing with ImageJ. Biophotonics International 11: 36–42.

[65] Moy TI, Ball AR, Anklesaria Z, Casadei G, Lewis K, Ausubel FM (2006) Identification of novel antimicrobials using a live-animal infection model. Proceedings of the National Academy of Sciences 103: 10414–10419, URL http://dx.doi.org/10.1073/pnas.0604055103.

[66] O'Rourke E, Conery A, Moy T (2009) Whole animal high-throughput screens: The *C. elegans* model. Methods Mol Biol 486: 57–75, URL http://dx.doi.org/10.1007/978-1-60327-545-3_5.

[67] Lee SS, Lee RYN, Fraser AG, Kamath RS, Ahringer J, Ruvkun G (2003) A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. Nature Genetics 33: 40–48, URL http://view.ncbi.nlm.nih.gov/pubmed/12447374.

[68] Hamilton B, Dong Y, Shindo M, Liu W, Odell I, Ruvkun G, Lee SS (2005) A systematic RNAi screen for longevity genes in *C. elegans*. Genes & Development 19: 1544–1555, URL http://dx.doi.org/10.1101/gad.1308205.

[69] Curran SPP, Ruvkun G (2007) Lifespan regulation by evolutionarily conserved genes essential for viability. PLoS Genet 3, URL http://dx.doi.org/10.1371/journal.pgen.0030056.

[70] Moy TI, Mylonakis E, Calderwood SB, Ausubel FM (2004) Cytotoxicity of hydrogen peroxide produced by *Enterococcus faecium*. Infect Immun 72: 4512–4520, URL http://view.ncbi.nlm.nih.gov/pubmed/15271910.

[71] Ashrafi K, Chang FY, Watts JL, Fraser AG, Kamath RS (2003) Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. Nature 421: 268–272, URL http://view.ncbi.nlm.nih.gov/pubmed/12529643.

[72] Mak HY, Nelson LS, Basson M, Johnson CD, Ruvkun G (2006) Polygenic control of *Caenorhabditis elegans* fat storage. Nature Genetics 38: 363–368, URL http://dx.doi.org/10.1038/ng1739.

[73] Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual JF, Kennedy S, Dybbs M, Bertin N, Kaplan JM, Vidal M, Ruvkun G (2005) Functional genomic analysis of RNA interference in *C. elegans*. Science 308: 1164–1167, URL http://dx.doi.org/10.1126/science.1109267.

[74] Breger J, Fuchs BB, Aperis G, Moy TI, Ausubel FM, Mylonakis E (2007) Antifungal chemical compounds identified using a *C. elegans* pathogenicity assay. PLoS Pathog 3: e18, URL http://dx.doi.org/10.1371/journal.ppat.0030018.

[75] Vidal M, Sieburth DS, Chng Q, Dybbs M, Tavazoie M, Kennedy S, Wang D, Dupuy D, Rual JF, Hill D (2005) Systematic analysis of genes required for synapse structure and function. In: International Worm Meeting, URL http://www.wormbase.org/db/misc/paper?name=WBPaper00026441.

[76] Miyata S, Begun J, Troemel ER, Ausubel FM (2008) DAF-16-dependent suppression of immunity during reproduction in *C. elegans*. Genetics URL http://view.ncbi.nlm.nih.gov/pubmed/18245330.

[77] Frand AR, Russel S, Ruvkun G (2005) Functional genomic analysis of *C. elegans* molting. PLoS Biol 3: e312, URL http://dx.doi.org/10.1371/journal.pbio.0030312.

[78] Parry DH, Xu J, Ruvkun G (2007) A whole-genome RNAi screen for *C. elegans* miRNA pathway genes. Curr Biol URL http://dx.doi.org/10.1016/j.cub.2007.10.058.

[79] Hu PJ, Xu J, Ruvkun G (2006) Two membrane-associated tyrosine phosphatase homologs potentiate *C. elegans* AKT-1/PKB signaling. PLoS Genetics 2: e99+, URL http://dx.doi.org/10.1371/journal.pgen.0020099.

[80] Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, Inoue H, Tanaka-Hino M, Hisamoto N, Matsumoto K, Tan MW, Ausubel FM (2002) A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. Science 297: 623–626, URL http://view.ncbi.nlm.nih.gov/pubmed/12142542.

[81] Paradis S, Ruvkun GB (1998) *Caenorhabditis elegans* Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. Genes and Development 12: 2488–2498, URL http://view.ncbi.nlm.nih.gov/pubmed/9716402.

[82] Dupuy D, Bertin N, Hidalgo C, Venkatesan K, Tu D, Lee D, Rosenberg J, Svrzikapa N, Blanc A, Carnec A, Carvunis AR, Pulak R, Shingles J, Reece-Hoyes J, Hunt-Newbury R, Viveiros R, Mohler W, Tasan M, Roth F, Le Peuch C, Hope I, Johnsen R, Moerman D, Barabasi AL, Baillie D, Vidal M (2007) Genome-scale analysis of in vivo spatiotemporal promoter activity in *Caenorhabditis elegans*. Nat Biotech 25: 663–668.

[83] Doitsidou M, Flames N, Lee AC, Boyanov A, Hobert O (2008) Automated screening for mutants affecting dopaminergic-neuron specification in *C. elegans*. Nat Meth 5: 869–872, URL http://dx.doi.org/10.1038/nmeth.1250.

[84] Mahajan-Miklos S, Tan MW, Rahme LG, Ausubel FM (1999) Molecular mechanisms of bacterial virulence eluciated using a *Pseudomonas aeruginosa–Caenorhabditis elegans* pathogenesis model. Cell 96: 47–56, URL http://view.ncbi.nlm.nih.gov/pubmed/9989496.

[85] Tan MW, Mahajan-Miklos S, Ausubel FM (1999) Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. PNAS 96: 715–720, URL http://www.pnas.org/cgi/content/abstract/96/2/715.

[86] Aballay A, Yorgey P, Ausubel FM (2000) *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of Caenorhabditis elegans. Current Biology 10: 1539–1542, URL http://view.ncbi.nlm.nih.gov/pubmed/11114525.

[87] Garsin DA, Sifri CD, Mylonakis E, Qin X, Singh KV, Murray BE, Calderwood SB, Ausubel FM (2001) A simple model host for identifying Gram-positive virulence factors. Proceedings of the National Academy of Sciences USA 98: 10892–10897, URL http://view.ncbi.nlm.nih.gov/pubmed/11535834.

[88] Sifri CD, Begun J, Ausubel FM, Calderwood SB (2003) *Caenorhabditis elegans* as a model host for *Staphylococcus aureus* pathogenesis. Infection and Immunity 71: 2208–2217, URL http://view.ncbi.nlm.nih.gov/pubmed/12654843.

[89] Mylonakis E, Ausubel FM, Perfect JR, Heitman J, Calderwood SB (2002) Killing of *Caenorhabditis elegans* by *Cryptococcus neoformans* as a model of yeast pathogenesis. Proceedings of the National Academy of Sciences USA 99: 15675–15680, URL http://view.ncbi.nlm.nih.gov/pubmed/12438649.

[90] Sifri C, Ausubel F (2005) Use of simple nonvertebrate hosts to model mammalian pathogenesis. In: Cossart P, editor, Cellular Microbiology (2nd edition), pp. 543–563, ASM Press.

[91] Tan MW, Rahme LG, Sternberg JA, Tompkins RG, Ausubel FM (1999) *Pseudomonas aeruginosa* killing of *Caenorhabditis elegans* used to identify *P. aeruginosa* virulence factors. Proceedings of the National Academy of Sciences USA 96: 2408–2413, URL http://view.ncbi.nlm.nih.gov/pubmed/10051655.

[92] Yorgey P, Rahme LG, Tan MW, Ausubel FM (2001) The roles of mucD and alginate in the virulence of *Pseudomonas aeruginose* in plants, nematodes and mice. Molecular Microbiology 41: 1063–1076, URL http://view.ncbi.nlm.nih.gov/pubmed/11555287.

[93] Garsin DA, Villanueva JM, Begun J, Kim DH, Sifri CD, Calderwood SB, Ruvkun G, Ausubel FM (2003) Long-lived *C. elegans* daf-2 mutants are resistant to bacterial pathogens. Science 300: 192–200, URL http://view.ncbi.nlm.nih.gov/pubmed/12817143.

[94] Kim DH, Ausubel FM (2005) Evolutionary perspectives on innate immunity from the study of *Caenorhabditis elegans*. Curr Opin Immunol 17: 4–10, URL http://view.ncbi.nlm.nih.gov/pubmed/15653303.

[95] Ausubel FM (2005) Are innate immune signaling pathways in plants and animals conserved? Nature Immunology 6: 973–979, URL http://dx.doi.org/10.1038/ni1253.

[96] Irazoqui JE, Ng A, Xavier RJ, Ausubel FM (2008) Role for beta-catenin and HOX transcription factors in *Caenorhabditis elegans* and mammalian host epithelial-pathogen interactions. Proceedings of the National Academy of Sciences 105: 17469–17474, URL http://dx.doi.org/10.1073/pnas.0809527105.

[97] Irazoqui JE, Urbach JM, Ausubel FM (2010) Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates. Nat Rev Immunol 10, URL http://dx.doi.org/10.1038/nri2689.

[98] Reinhart B, Slack F, Basson M, Pasquinelli A, Bettinger J, Rougvie A, Horvitz H, Ruvkun G (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403: 901–6.

[99] Pasquinelli A, Reinhart B, Slack F, Martindale M, Kuroda M, Maller B, Hayward D, Ball E, Degnan B, Mller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408: 86–9.

[100] Ogg S, Paradis S, Gottlieb S, Patterson G, Lee L, Tissenbaum H, Ruvkun G (1997) The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. Nature 389: 994–9.

[101] Wählby C, Lindblad J, Vondrus M, Bengtsson E, Björkesten L (2002) Algorithms for cytoplasm segmentation of fluorescence labelled cells. Anal Cell Patho 24: 101–11.

[102] Lindblad J, Wählby C, Bengtsson E, Zaltsman A (2004) Image analysis for automatic segmentation of cytoplasms and classification of Rac1 activation. Cytometry A 57: 22–33.

[103] Wählby C, Sintorn IM, Erlandsson F, Borgefors G, Bengtsson E (2004) Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. J Microscopy 215: 67–76, URL http://dx.doi.org/10.1111/j.0022-2720.2004.01338.x.

[104] Erlandsson F, Linnman-Wählby C, Ekholm S, Bengtsson E, Zetterberg A (2000) A detailed analysis of cyclin A accumulation at the G(1)/S border in normal and transformed cells. Exp Cell Res 259: 86–95.

[105] Erlandsson F, Wählby C, Ekholm-Reed S, Hellström A, Bengtsson E, A Z (2003) Abnormal expression pattern of cyclin E in tumour cells. Int J Cancer 104: 369–75.

[106] Jahangir Tafrechi R, van de Rijke F, Allalou A, Larsson C, Sloos W, van de Sande M, Wählby C, Janssen G, Raap A (2007) Single-cell A3243G mitochondrial DNA mutation load assays for segregation analysis. J Histochem Cytochem 55: 1159–66.

[107] Pinidiyaarachchi A, Zieba A, Allalou A, Pardali K, Wählby C (2009) A detailed analysis of 3D subcellular signal localization. Cytometry A 75: 319–28.

[108] Jarvius M, Paulsson J, Weibrecht I, Leuchowius K, Andersson A, Wählby C, Gullberg M, Botling J, Sjöblom T, Markova B, Ostman A, Landegren U, Söderberg O (2007) In situ detection of phosphorylated platelet-derived growth factor receptor beta using a generalized proximity ligation method. Mol Cell Proteomics 6: 1500–9.

[109] Allalou A, Wählby C (2009) Blobfinder, a tool for fluorescence microscopy image cytometry. Comput Methods Programs Biomed 94: 58–65.

[110] Allalou A, Pinidiyaarachchi A, Wählby C (2010) Robust signal detection in 3D fluorescence microscopy. Cytometry A 77: 86–96.

[111] Gavrilovic M, Wählby C (2009) Quantification of colocalization and cross-talk based on spectral angles. J Microsc 234: 311–24.

[112] Göransson J, Wählby C, Isaksson M, Howell W, Jarvius J, Nilsson M (2009) A single molecule array for digital targeted molecular analyses. Nucleic Acids Res 37: e7.

[113] Wählby C, Erlandsson F, Bengtsson E, Zetterberg A (2002) Sequential immunofluorescence staining and image analysis for detection of large numbers of antigens in individual cell nuclei. Cytometry 47: 32–41.

[114] Zieba A, Wählby C, Hjelm F, Jordan L, Berg J, Landegren U, Pardali K (2010) Bright-field microscopy visualization of proteins and protein complexes by in situ proximity ligation with peroxidase detection. Clin Chem 56: 99–110.

[115] Jones TR (2007) Predicting Gene Function from Images of Cells. Ph.D. thesis, Massachusetts Institute of Technology.

[116] Jones T, Carpenter A, Sabatini D, Golland P (2008) Methods for high-content, high-throughput–based cell screening. In: Rittscher J, Wong STC, Machiraju R, editors, Microscopic Image Analysis for Life Science Applications, pp. 209–221, Artech House.

[117] Wheeler D, Bailey S, Guertin D, Carpenter AE, Higgins C, Sabatini DM (2004) Microarrays of lentiviruses for gene function screens in immortalized and primary cells. Nature Methods 1: 127–132, URL http://dx.doi.org/10.1038/nmeth711.

[118] Cowen LE, Lindquist S (2005) Hsp90 potentiates the rapid evolution of new traits: Drug resistance in diverse fungi. Science 309: 2185–2189, URL http://dx.doi.org/10.1126/science.1118370.

[119] Cowen LE, Carpenter AE, Matangkasombut O, Fink GR, Lindquist S (2006) Genetic architecture of Hsp90-dependent drug resistance. Eukaryotic Cell 5: 2184–2188, URL http://dx.doi.org/10.1128/EC.00274-06.

[120] Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, Carpenter AE, Foo SY, Stewart SA, Stockwell BR, Hacohen N, Hahn WC, Lander ES, Sabatini DM, Root DE (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124: 1283–1298, URL http://dx.doi.org/10.1016/j.cell.2006.01.040.

[121] Baltus AE, Menke DB, Hu YC, Goodheart ML, Carpenter AE, de Rooij DG, Page DC (2006) In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nature Genetics 38: 1430–1434, URL http://dx.doi.org/10.1038/ng1919.

[122] Sigal A, Milo R, Cohen A, Geva-Zatorsky N, Klein Y, Alaluf I, Swerdlin N, Perzov N, Danon T, Liron Y, Raveh T, Carpenter AE, Lahav G, Alon U (2006) Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. Nat Meth 3: 525–531, URL http://dx.doi.org/10.1038/nmeth892.

[123] Bailey SN, Ali SM, Carpenter AE, Higgins CO, Sabatini DM (2006) Microarrays of lentiviruses for gene function screens in immortalized and primary cells. Nature Methods 3: 117–122, URL http://dx.doi.org/10.1038/nmeth848.

[124] Hartwell KA, Muir B, Reinhardt F, Carpenter AE, Sgroi DC, Weinberg RA (2006) The Spemann organizer gene, Goosecoid, promotes tumor metastasis. Proceedings of the National Academy of Sciences 103: 18969–18974, URL http://dx.doi.org/10.1073/pnas.0608636103.

[125] Lamprecht MR, Sabatini DM, Carpenter AE (2007) CellProfiler: Free, versatile software for automated biological image analysis. Biotechniques 42: 71–75.

[126] Pan H, Yu J, Zhang L, Carpenter A, Zhu H, Li L, Ma D, Yuan J (2008) A novel small molecule regulator of guanine nucleotide exchange activity of ARF and golgi membrane trafficking. J Biol Chem pp. M806592200+, PMCID PMC2576541, URL http://dx.doi.org/10.1074/jbc.M806592200.

[127] Tieu K, Viola P (2004) Boosting image retrieval. International Journal of Computer Vision 56: 17–36.

[128] Tsui M, Xie T, Orth J, Carpenter A, Rudnicki S, Kim S, Shamu C, Mitchison T (2009) An intermittent live cell imaging screen for siRNA enhancers and suppressors of a kinesin-5 inhibitor. PLoS One 4: e7339, PMCID PMC2752188, URL http://dx.doi.org/10.1371/journal.pone.0007339.

[129] Golland P, Grimson W, Kikinis R (1999) Statistical shape analysis using fixed topology skeletons: Corpus callosum study. In: In Proceedings of IPMI: International Conference on Information Processing and Medical Imaging, LNCS, vol. 1613, pp. 382–387.

[130] Golland P, Grimson W, Shenton M, Kikinis R (2001) Small sample size learning for shape analysis of anatomical structures. In: Proceedings of the International Conference on Medical Computing and Computer Assisted Intervention, vol. 1935, pp. 72–82.

[131] Golland P, Grimson W, Shenton M, Kikinis R (2001) Deformation analysis for shaped based classification. In: In Proceedings of IPMI: International Conference on Information Processing and Medical Imaging, LNCS, vol. 2082, pp. 517–530.

[132] Golland P, Grimson W, Shenton M, Kikinis R (2000) Fixed topology skeletons. In: Proc. CVPR, pp. 10–17.

[133] Golland P (2001) Discriminative direction for kernel classifiers. In: NIPS, pp. 745–752.

[134] Van Leemput K, Bakkour A, Benner T, Wiggins G, Wald L, Augustinack J, Dickerson B, Golland P, Fischl B (2008) Model-based segmentation of hippocampal subfields in ultra-high resolution in vivo MRI. In: Proceedings of the International Conference on Medical Computing and Computer Assisted Intervention, accepted.

[135] Riklin-Raviv T, Kiryati N, Sochen N (2004) Unlevel-sets: Geometry and prior-based segmentation. In: Proceedings of the European Conference on Computer Vision, vol. 4, pp. 50–61.

[136] Riklin-Raviv T, Kiryati N, Sochen N (2005) Prior-based segmentation by projective registration and level-sets. In: Proceedings of the Workshop on Computer Vision for Biomedical Image Applications at the International Conference of Computer Vision (ICCV), pp. 204–211.

[137] Riklin-Raviv T, Kiryati N, Sochen N (2006) Segmentation by level-sets and symmetry. In: Proc. CVPR, vol. 1, pp. 1015–1022.

[138] Riklin-Raviv T, Sochen N Nand Kiryati (2006) Mutual segmentation with level-sets. In: The fifth IEEE computer society workshop on perceptual organization in computer vision.

[139] Riklin-Raviv T, Kiryati N, Sochen N (2007) Prior-based segmentation and shape registration in the presence of perspective distortion. International Journal of Computer Vision 72: 309–328, URL http://dx.doi.org/10.1007/s11263-006-9042-y.

[140] Riklin-Raviv T, Sochen N, Kiryati N (2007) Shape-based mutual segmentation. International Journal of Computer Vision Online Publication, DOI: 10.1007/s11263-007-0115-3.

[141] Riklin-Raviv T, Sochen N, Kiryati N (2008) On symmetry, projectivity and level-set based segmentation. IEEE Trans PAMI Accepted.

[142] Vokes MS, Carpenter AE (2008) Using CellProfiler for automatic identification and measurement of biological objects in images. Curr Protoc Mol Biol URL http://dx.doi.org/10.1002/0471142727.mb1417s82.

[143] Jones T, Carpenter A (2007) Flexible image analysis. Biophotonics International 14: 31–32.

[144] Carpenter A (2009) Data analysis: Extracting rich information. Methods Mol Biol 486.

[145] Kuijl C, Savage NDL, Marsman M, Tuin AW, Janssen L, Egan DA, Ketema M, van den Nieuwendijk R, van den Eeden SJF, Geluk A, Poot A, van der Marel G, Beijersbergen RL, Overkleeft H, Ottenhoff THM, Neefjes J (2007) Intracellular bacterial growth is controlled by a kinase network around PKB/AKT1. Nature 450: 725–730, URL http://dx.doi.org/10.1038/nature06345.

[146] Wiley DJ, Rosqvist R, Schesser K (2007) Induction of the *Yersinia* type 3 secretion system as an all-or-none phenomenon. Journal of Molecular Biology 373: 27–37, URL http://dx.doi.org/10.1016/j.jmb.2007.07.077.

[147] Lian S, Fritzler MJ, Katz J, Hamazaki T, Terada N, Satoh M, Chan EK (2007) Small interfering RNA-mediated silencing induces target-dependent assembly of GW/P bodies. Mol Biol Cell 18: 3375–3387, URL http://dx.doi.org/10.1091/mbc.E07-01-0070.

[148] Mu JJ, Wang Y, Luo H, Leng M, Zhang J, Yang T, Besusso D, Jung SY, Qin J (2007) A proteomic analysis of ataxia telangiectasia-mutated (ATM)/ATM-Rad3-related (ATR) substrates identifies the ubiquitin-proteasome system as a regulator for dna damage checkpoints. J Biol Chem 282: 17330–17334, URL http://dx.doi.org/10.1074/jbc.C700079200.

[149] Schlumberg M, Käpeli R, Wetter M, Miller AJ, Misselwitz B, Dilling S, Kremer M, Hardt WD (2007) Two newly identified SipA domains (F1, F2) steer effector protein localization and contribute to *Salmonella* host cell manipulation. Molecular Microbiology 65: 741–60, URL http://dx.doi.org/10.1111/j.1365-2958.2007.05823.x.

[150] Dai C, Whitesell L, Rogers AB, Lindquist S (2007) Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. Cell 130: 1005–1018, URL http://dx.doi.org/10.1016/j.cell.2007.07.020.

[151] Dobbelaere J, Josué F, Suijkerbuijk S, Baum B, Tapon N, Raff J (2008) A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in drosophila. PLoS Biology 6, URL http://dx.doi.org/10.1371/journal.pbio.0060224.

[152] Fjeldbo CSS, Misund K, Günther CCC, Langaas M, Steigedal TSS, Thommesen L, Lægreid A, Bruland T (2008) Functional studies on transfected cell microarray analysed by linear regression modelling. Nucleic acids research URL http://view.ncbi.nlm.nih.gov/pubmed/18628295.

[153] Zunder ER, Knight ZA, Houseman BT, Apsel B, Shokat KM (2008) Discovery of drug-resistant and drug-sensitizing mutations in the oncogenic PI3K isoform p110α. Cancer Cell 14: 180–192, URL http://dx.doi.org/10.1016/j.ccr.2008.06.014.

[154] Ramo P, Sacher R, Snijder B, Begemann B, Pelkmans L (2009) CellClassifier: supervised learning of cellular phenotypes. Bioinformatics 25: 3028–3030, URL http://dx.doi.org/10.1093/bioinformatics/btp524.

[155] Snijder B, Sacher R, Ramo P, Damm EM, Liberali P, Pelkmans L (2009) Population context determines cell-to-cell variability in endocytosis and virus infection. Nature 461: 520–523, URL http://dx.doi.org/10.1038/nature08282.

[156] Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models—their training and application. Computer Vision and Image Understanding 61: 38–59, URL http://www.sciencedirect.com/science/article/B6WCX-45NJT56-1K/2/1933a32b2a89e4c3f032ba0718f4ef03.

[157] Blum H (1967) A transformation for extracting new descriptors of shape. In: Wathen-Dunn W, editor, Models for the Perception of Speech and Visual Form: Proceedings of a Symposium, pp. 362–380, MIT Press.

[158] Golland P, Grimson W, Shenton M, Kikinis R (2001) Deformation analysis for shape based classification. In: IPMI, pp. 517–530.

[159] Styner M, Gerig G, Joshi SC, Pizer SM (2003) Automatic and robust computation of 3D medial models incorporating object variability. International Journal of Computer Vision 55: 107–122.

[160] Goutsias J, Shonfeld D (1991) Morphological representation of discrete and binary images. IEEE Trans Signal Proc 39: 1369–1379, URL http://dx.doi.org/10.1109/78.136543.

[161] Johnson R, Wichern DW (1988) Multivariate Statistics, A Practical Approach. Chapman & Hall.

[162] Leventon ME, Grimson WEL, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proc IEEE Conf Comp Vis and Pattern Recog, vol. 1, pp. 316–323, URL http://dx.doi.org/10.1109/CVPR.2000.855835.

[163] Cormen TH, Leiserson CE, abd Clifford Stein RLR (2001) Introduction to Algorithms. Addison-Wesley.

[164] Geng W, Cosman P, Berry CC, Feng Z, Schafer WR (2004) Automatic tracking, feature extraction and classification of *C. elegans* phenotypes. IEEE Transactions on Biomedical Engineering 51: 1811–1820, URL http://dx.doi.org/10.1109/TBME.2004.831532.

[165] Zhang J, Chung T, Oldenburg K (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen 4: 67–73.

[166] Vincent L, Soille P (1991) Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans PAMI 13: 583–598.

[167] Gonzales RC, Woods RE (1993) Digital Image Processing. Addison-Wesley.

[168] Lindblad J, Bengtsson E (2001) A comparison of methods for estimation of intensity nonuniformities in 2D and 3D microscope images of fluorescence stained cells. In: Proc Scandinavian Conf Image Analysis (SCIA), pp. 264–271.

[169] Osher S, Sethian J (1988) Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics 79: 12–49.

[170] Chan T, Vese L (2001) Active contours without edges. IEEE Transactions on Image Processing 10: 266–277.

[171] Paragios N, Deriche R (2002) Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision. Journal of Visual Communication and Image Representation 13: 249–268.

[172] Kichenassamy S, Kumar A, Olver P, Tannenbaum A, Yezzi A (1995) Gradient flows and geometric active contour models. In: Proceedings of the Workshop on Computer Vision for Biomedical Image Applications at the International Conference of Computer Vision (ICCV), pp. 810–815.

[173] Caselles V, Kimmel R, Sapiro G (1997) Geodesic active contours. International Journal of Computer Vision 22: 61–79.

[174] Vasilevskiy A, Siddiqi K (2001) Flux maximizing geometric flows. In: Proceedings of the Workshop on Computer Vision for Biomedical Image Applications at the International Conference of Computer Vision (ICCV), pp. 149–154.

[175] Kimmel R, Bruckstein A (2003) Regularized Laplacian zero crossings as optimal edge integrators. International Journal of Computer Vision 53: 225–243.

[176] Serra J (1989) Image Analysis and Mathematical Morphology, vol. 1. Academic Press, London.

[177] Cremers D, Kohlberger T, Schnörr C (2003) Shape statistics in kernel space for variational image segmentation. Pattern Recognition 36: 1929–1943.

[178] Cremers D, Rousson M, Deriche R (2007) A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. International Journal of Computer Vision 72: 195–215.

[179] Chen Y, Tagare H, Thiruvenkadam S, Huang F, Wilson D, Gopinath K, Briggs R, Geiser E (2002) Using prior shapes in geometric active contours in a variational framework. International Journal of Computer Vision 50: 315–328.

[180] Leventon M, Grimson W, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proc. CVPR, vol. I, pp. 316–323.

[181] Rousson M, Paragios N (2002) Shape priors for level set representation. In: Proceedings of the European Conference on Computer Vision, pp. 78–92.

[182] Long F, Peng H, Liu X, Kim SK, Myers E (2009) A 3D digital atlas of *C. elegans* and its application to single-cell analyses. Nat Methods 6: 667–672, URL http://dx.doi.org/10.1038/nmeth.1366.

[183] Torralba A, Murphy K, Freeman W (2007) Sharing visual features for multiclass and multiview object detection. IEEE Trans PAMI 29: 854–869, URL http://dx.doi.org/10.1109/TPAMI.2007.1055.

MASSACHUSETTS GENERAL HOSPITAL

HARVARD MEDICAL SCHOOL

Department of Molecular Biology
Richard B. Simches Research Center
185 Cambridge Street, 7th Floor
Boston, MA 02114
Tel: 617.726.5969, Fax: 617.726.5949
E-mail: ausubel@molbio.mgh.harvard.edu

Frederick M. Ausubel, Ph.D.
Professor of Genetics
Harvard Medical School
Massachusetts General Hospital

January 22, 2010

Carolina Wählby
Broad Institute of Harvard and MIT
Uppsala University

Dear Carolina,

I am eager to continue our successful collaboration as part of your NIH R01 proposal, “Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism”. Our work together over the past couple of years, which has been built on the foundation of our long-term collaboration with Anne Carpenter’s group at the Broad Institute, has demonstrated that whole-organism screens using adult *C. elegans* in 384-well plates are feasible and effective for identifying chemicals that affect the course of a bacterial infection. Automating image analysis for distinguishing live and dead *C. elegans* by shape and fluorescent staining is revolutionizing our ability to rapidly screen large chemical compound libraries and RNAi libraries more accurately and rapidly than has been possible, as well as enabling screens that cannot be accomplished with current technology. I have a great deal of confidence in your ability to manage this large-scale project with its various contributing personnel.

My research group has pioneered the development of so-called multi-host pathogenesis systems that involve the infection of invertebrate hosts with human bacterial pathogens. Together with Gary Ruvkun, we have also pioneered and assembled an automated *C. elegans* sample preparation pipeline over the past two years to enable high-throughput image-based screens. We used the system to screen 37,000 chemicals for regulators of infection of *C. elegans* by the important human opportunistic pathogen *Enterococcus faecalis*, the first screen of adult *C. elegans* to employ automated image analysis (Moy, T.I., A.L. Connery, J. Larkins-Ford, G. Wu, R. Mazitschek, G. Casadei, K. Lewis, A.E. Carpenter and F.M. Ausubel (2009) High throughput screen for antimicrobials using a whole animal infection model. *ACS Chem. Biol.,* 4:527-533). The screen was highly successful, leading to the identification of at least three new chemical backbones with antimicrobial activities that are amenable to medicinal chemistry analysis. In the Moy et al. paper (cited above), live and dead worms were distinguished by staining dead worms with the fluorescent dye Sytox Orange. However, our recent joint paper that was recently accepted for publication, substantially advances the image analysis technology and sets the stage for developing assays that do not depend on staining dead worms with fluorescent dyes (Wählby, C., T. Riklin-Raviv, V. Ljosa, A.L. Conery, P. Golland, F.M. Ausubel and A.E. Carpenter (2010) Resolving clustered worms via probabilistic shape models. Accepted for publication in proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI) 2010).

Now that we have successfully assembled the hardware and worked out the details of the liquid killing assays, we will be able to work with you to develop more sophisticated image analysis

software that will enable us to carry out much larger screens. In addition to the images that we have already collected for the *E. faecalis* screen, we also have collected control images for the other assays. All of these images are available for your work. The *E. faecalis* screen has laid the groundwork for additional full-scale chemical screens employing new assays that we have developed, which are funded through other grants in my laboratory:

1. A live/dead assay to identify modulators of infection of *C. elegans* by *microsporidia*, obligate intracellular eukaryotic parasites that are an important emerging human pathogen and that has recently been added to the NIH list of priority pathogens.
2. Live/dead assay to identify compounds that cure a *Staphylococcus aureus* or a *Pseudomonas aeruginosa* infection of the *C. elegans* intestine.
3. A protein localization assay to identify modulators of the GFP-labeled proteins *clec-60* and/or *irg-1::GFP* in response to infection of *C. elegans* by *S. aureus* or *P. aeruginosa*, respectively. These GFP reporters change their pattern of expression in the *C. elegans* intestine in response to infection and this screen will identify potential regulators of location-specific immunity pathways.

The funding that I have for these projects covers the development of *C. elegans* assays. Therefore, my group will be able to assist you in developing automated image analysis for the assays described above. Specifically, we will be able to provide:

1. Expert annotation of control sets of images for the new assays.
2. Informatics support to assist transfer of images and annotations to your group.
3. Collaboration to assess the results of automated scoring.

To guide this work, my group members working on the project will attend at least a monthly project meeting with your researchers, and I will join them at a twice-yearly meeting of the entire project team, including PI's. The scientists in my group who have developed the assays and the high-throughput automated methodology for sample preparation and image acquisition will be committed to contributing to the project. These include Dr. Annie Lee-Conery, the project manager and Mr. Jonah Ford-Larkins, a technician in my laboratory who has been the primary developer of many of the automated robotic steps of the screening assays. As you know, we have proven quite proficient at learning the methodology and software that your group produces and are looking forward to next-generation algorithms to solve these more complex assays.

We will continue to work closely with you as you optimize the computational methods for analyzing images for this important work. I wish you the best with your grant application.

Sincerely,

Frederick M. Ausubel

BROAD INSTITUTE

7 Cambridge Center
Cambridge, MA 02142
T 617-714-7000 F 617-714-8972
www.broadinstitute.org

Dr. Carolina Wählby
Broad Institute & Uppsala University

January 22, 2010

Dear Carolina,

With great pleasure, I write to support your NIH R01 proposal, "Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism".

As you know, I became interested in developing image analysis algorithms for *C. elegans* through collaborations that began with Fred Ausubel and Gary Ruvkun in 2004. While helping them to build the *C. elegans* Screening Center at Harvard/MGH and to identify novel regulators of infection through this collaboration, it became clear there is a huge need for image analysis for this organism, which enables screening in the physiological context of a whole animal. However, my group has been unable to be devote the effort necessary to "solve" the many assays that *C. elegans* researchers have since brought to us which would otherwise be ready-to-screen; my Imaging Platform has focused on cell-based assays.

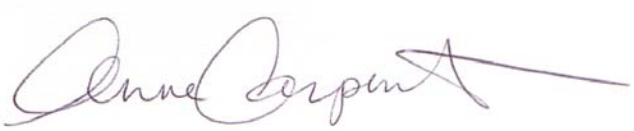
I was therefore thrilled to learn that you decided to direct the computational expertise of your group to this long-standing need in the field upon joining the Broad Institute. You have already made tremendous progress since being introduced to the *C. elegans* researchers and the challenges my group had encountered.

Based on my own group's experience developing computational methods to extract information from microscopy images and helping biologists to apply the resulting software, we are well equipped to provide assistance for the work you propose. My team will assist in the validation of the algorithms your team develops, but primarily will focus on implementing the algorithms into our user-friendly open-source software (CellProfiler) and disseminating the result to the *C. elegans* community. In particular, we will organize workshops and tutorials to teach the software to those interested in image-based screening in *C. elegans,* building on our more general workshops.

I have every confidence that your strong image processing background, combined with the expertise of my group and the other collaborators, will produce a validated toolbox of algorithms for *C. elegans* high-throughput assays and thus substantially transform the field.

I therefore look forward to a continued productive collaboration with you and your group.

Sincerely,



Anne Carpenter, Ph.D.
Director, Imaging Platform
Co-founder, CellProfiler project
Broad Institute of MIT and Harvard

OLINK
BIOSCIENCE

Uppsala January 26, 2010

Centre for Scientific Review
National Institutes for Health
6701 Rockledge Drive
Room 1040
MSC 7710
Bethesda, MD 20817

Letter of Support

With this letter we wish to declare our enthusiastic support for Dr. Carolina Wählby's R01 proposal "Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism".

We have been working with Dr. Wählby in an EU funded three-year project called ENLIGHT (ENhanced LIGase based Histochemical Techniques), starting in mid 2006. The project had the purpose to develop highly specific analytical procedures to analyse individual molecular events in fixed cells and tissue section (*in situ*). Detected targets included nucleic acids and proteins and also interacting proteins and protein modifications. Reagents and methods for the analysis of specific biomarkers of particular interest in oncology were developed. The padlock technology was developed for interrogation of nucleic acids, and the proximity ligation assay (PLA) for protein analysis.

The project was organised in three work packages (WPs). One of the WP's addressed the development of new image analyses tools required for single molecular imaging applications including automatic identification, counting and localization of PLA and padlock signals. Dr Wählby was professionally managing this WP and developed a most productive collaboration with the other partner in the imaging WP, VisioPharm A/S, a Danish image analysis software company. Under the management of Dr Wählby the WP was efficiently organized and highly productive and fulfilled all obligations defined in the project plan. Besides the effective work within the WP, she also contributed significantly to productive interactions with the other WPs including researchers from other scientific disciplines (e.g. biology, medicine).

Dr Wählby showed great scientific skills and excellent leadership, and clearly contributed to the success of the project. Dr Wählby was engaged, energetic and hard-working but also social and appreciated as a person to interact with. We are convinced that Dr Wählby would do an excellent job in the proposed R01 project. Please let us know if you need any further information.

Björn Ekström
President and CEO Olink AB
Coordinator of EU project ENLIGHT
Uppsala, SWEDEN
Phone: +46-706-554110
E-mail: bjorn.ekstrom@olink.com

Anders Alderborn
Former Project coordinator at Olink AB
Currently Project coordinator at Uppsala University
Coordinator of EU project ENLIGHT
Uppsala, SWEDEN
Phone: +46-706-585915
E-mail: anders.alderborn@genpat.uu.se

UPPSALA
UNIVERSITET

Ewert Bengtsson

Professor, head

Centre for Image Analysis
Centrum för bildanalys, CBA

Box 337
SE-75105 Uppsala, Sweden

Visiting address:
Lägerhyddvägen 2
Uppsala

Telefon/Phone:
018-471 3467
+46 18 471 3467

Telefax/Fax:
018-55 34 47
+46 18 55 34 47

www.cb.uu.se

ewert@cb.uu.se

LETTER OF SUPPORT

Uppsala, January 12th 2010

Center for Scientific Review
National Institutes of Health
6701 Rockledge Drive,
Room 1040, MSC 7710
Bethesda, MD 20817

Dear Sir/Madam,

It is with enthusiasm I write this letter of support for Dr Carolina Wählby´s R01 application "Image analysis algorithms for high-throughput C. elegans assays to identify regulators of infection and metabolism". I have had the pleasure of knowing Dr Wählby for quite some time now. She applied for a PhD position on "3D biomedical image analysis" in my group in 1998 and since she was clearly the most qualified applicant, already with a journal publication from her undergraduate work, she got the position. During her PhD work she did outstanding work, developing new approaches to cell image segmentation in 2D and 3D which led to several conference and journal papers. During her PhD period she collaborated quite independently from me with several groups in Sweden and abroad, and she spent some time as visiting researcher e.g. in Korea. Her PhD thesis in 2003 was one of the very best among the more than 25 PhD theses I have supervised over the years.

After her PhD she got a position as assistant professor at our centre in tough competition with several other well qualified applicants. In that capacity she continued to develop as an independent researcher of strong international standing. Through her we built a new close collaboration with professor Ulf Landegrens outstanding research group at the Rudbecks laboratory at Uppsala University where she took up a part time position as their image analysis expert in parallel to her position with us. She also became the leader for the work package on image analysis in the project ENLIGHT, Enhanced ligase-based histochemical techniques, which was funded by the European Union with about 3 million Euro from 2006 through 2009. The project involved research groups from several different countries and different disciplines so successfully managing a work-package in this context was challenging. Still she managed the image analysis part of the project so well that the image analysis results clearly exceeded the specifications in the original application. The project led to seven reviewed journal publications with Carolina as senior author on four of them plus

Organisationsnr:
202100-2932

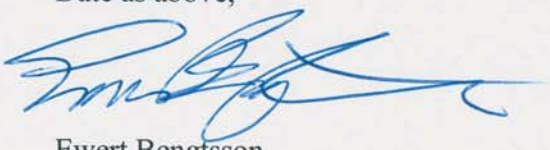
a reviewed conference article with her a senior author. In addition to these publications, the image analysis tools developed within the project was made available as freeware "BlobFinder" that can be downloaded from our home page and is being used by numerous groups world-wide. A new approach to quantification of co-localization and elimination of cross-talk in fluorescence microscopy images inspired by the research in the project was developed further and led to a patent application with Carolina as one of the inventors and one of the PhD students she is supervising another inventor.

She also took up collaboration with professor Nils-Göran Larsson at the Karolinska Institute and filed an application to the Swedish research Council jointly with his group for developing methods for studying mitochondrial DNA segregation. It was granted in spite of very strong competition and is currently supporting a PhD student here at CBA whom Carolina is supervising from the US, together with assistant local supervisors here.

Carolina has also been main supervisor for one student who got her PhD last year and assistant supervisor for two more PhD students and also supervisor for several Master thesis students. She became associate professor (docent) a year ago.

So in summary Carolina has shown impressive capabilities in conducting research and managing research projects. In particular she has shown great strength in creating fruitful collaborations between technical and biological researchers. She has done all this while building a family, giving birth to three children, taking an active part in developing our teaching curricula, maintaining our website, editing the national Swedish newsletter for the image analysis scientific community and many other tasks. I have no doubt whatsoever that she would do an excellent job if she is granted the funds she is now applying for.

Date as above,



Ewert Bengtsson
professor
Head of Centre for Image Analysis, CBA
Uppsala University, Sweden

MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

January 25, 2010

Carolina Wahlby
Broad Institute of Harvard and MIT
Uppsala University
7 Cambridge Center
Cambridge, MA 02142

Dear Carolina,

I am writing to express enthusiastic support for your grant proposal “Image analysis algorithms for high-throughput C. elegans assays to identify regulators of infection and metabolism.” The aims of this proposal represent a natural evolution of our recent joint work on automatic methods for microscopy image analysis in C. elegans experiments. I look forward to our continuing collaboration on this truly interdisciplinary project. My research focuses on computational modeling of biological shape, its relationship with biological function and its statistical variability. Shape detection and analysis from fluorescence microscopy images is an interesting challenge from the computational point of view. Segmentation and shape characterization of C. elegans provides a great application for computational and statistical models of biological shape.

This project builds on a long-term collaboration between my group and the Imaging Platform at the Broad Institute. As part of this collaboration, we contributed significantly to development of CellProfiler, a software system for high throughput analysis of cell images. I look forward to collaborating with your group on developing the next generation of algorithms for segmentation and shape analysis of C. elegans. I also enthusiastically support your plans to distribute the algorithms and the test image data sets to the community. I believe that open source software development and data sharing enable much faster scientific progress by engaging more researchers in development and validation of image analysis methods. I will be happy to see the algorithms and the data created in this project shared with the broad scientific community.

As part of this collaboration, my group will work closely with the researchers from your group to develop algorithms for segmentation and shape characterization of C. elegans in high throughput microscopy images. This problem is an instance of model-based segmentation and shape analysis, which is the area of expertise of several members of my group. Specifically, Tamar Riklin Raviv is an expert on implicit representations of shapes in images, shape priors and shape-based segmentation. Her knowledge in this field is directly applicable to the problems we face in microscopy image analysis. My expertise relevant to this project is in modeling shape distributions and characterizing statistical differences in shape across populations. Our initial joint efforts to solve this problem led to two

The Stata Center, Building 32-D470
32 Vassar Street, Cambridge, Massachusetts 02139-4307
PHONE +1.617.253.8005 FAX +1.617.253.4640
WEB csail.mit.edu

conference papers, and I expect many interesting publications to come out of this work in the future. I expect your group to lead the algorithm development and testing, with our help and guidance on model-based segmentation. I believe our prior experience in the field of shape modeling, detection and analysis will be directly applicable to this problem.

I expect our groups to meet at least monthly to coordinate our effort, with more extensive semiannual status update meetings. Your researchers are also welcome to continue attending the general scientific meetings in my group, as they have been doing for the past several years. This close community should provide a good venue for brainstorming and testing new ideas.

In addition to the biological benefits of this work, we anticipate that the algorithms developed in the course of this collaboration will in turn be useful for the biomedical image analysis domain, which is the primary focus of our work. Having seen the many benefits of sharing algorithms and ideas between our groups in the past, I am eager to collaborate with you on this important and challenging problem.

Sincerely,

Polina Golland

Polina Golland
Associate Professor
MIT EECS/CSAIL

MASSACHUSETTS GENERAL HOSPITAL

Eleftherios Mylonakis, M.D.
Associate Professor of Medicine
E-mail: emylonakis@partners.org

21 January 2010

Carolina Wahlby, Ph.D.
Broad Institute of Harvard and MIT

Dear Carolina,

I am writing to enthusiastically support your proposal, "Image analysis algorithms for high-throughput C. elegans assays to identify regulators of infection and metabolism" for NIH R01 funding.

I became aware of your work through Fred Ausubel's group. The recent publication of the first automatically-analyzed image-based screen *C. elegans*, a collaboration between his group and members of your project team, was certainly exciting for my group to see, given that we also are scaling up our use of this model organism for high-throughput screening.

My group is interested in the development of new antifungal therapies. We use *C. elegans* in order to identify novel antifungal compounds in a whole animal high assay. For this work, we find screening in a whole organism to be indispensible for probing the phenotypes of interest/studying this critical disease area.

Although we are not directly collaborating with you in this proposal, it is clear that the flexible, open-source software that you will develop will be directly applicable to multiple screens planned in my laboratory. I will gladly share images that we collect in our group for various assays so that your algorithms can be designed to be broadly applicable to other laboratories and assays.

I wish you the best in obtaining funding for your work and look forward to making productive use of the software tools that result!

Sincerely,

E. Mylonakis, M.D., Ph.D., FIDSA

Molecular Genetics
UNIVERSITY OF TORONTO
Donnelly Centre
for Cellular + Biomolecular Research

FACULTY OF MEDICINE, DEPARTMENT OF MOLECULAR GENETICS
THE TERRENCE DONNELLY CENTRE FOR CELLULAR & BIOMOLECULAR RESEARCH
Dr. Peter J. Roy, Associate Professor

January 22nd, 2010

Carolina Wahlby, Ph.D.
Broad Institute of Harvard and MIT
Uppsala University

Dear Carolina,

I am pleased to support your application to the NIH, entitled, "Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism". This work would benefit my lab and the many others who are interested in measuring properties of this important model organism to answer a wide variety of biological questions in a whole, living animal.

As you know, my research group has ongoing interests in identifying and characterizing new small molecule structures that perturb conserved regulators of animal development with the ultimate goal of generating new molecular tools for biological investigation and potential drug leads. Our basic approach is to co-cultivate C. elegans with the small molecule in 24-well plate format and then examine progeny 3-6 days later for phenotypes consistent with defects in development. We have screened over 70,000 small molecules in worms or yeast (as a prescreen) and have over 1000 bioactive molecules that have peaked our interest. We are currently investigating novel antagonists of the Wnt and insulin signaling pathways in great detail in a similar fashion as we have done previously with a calcium channel antagonist (Kwok *et al*., *Nature*, 2006).

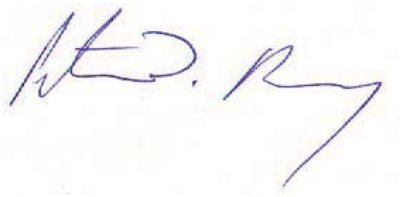
To date, we have mainly analyzed the experimental samples from our screens manually, and have archived most of our experiments using an automated imager that we developed (described in Burns, *et al*., *Nature Protocols*, 2006). Although we have developed image analysis software to measure certain properties of the images (*ibid*), the algorithm does not distinguish individual small worms (L1 and L2 larvae) and can encounter difficulties when larger worms are especially dense. I would be happy to continue to share any or all of these 1000s of archived images with you to help guide your algorithm development for a robust image analysis software package. I was pleased to hear that the images our group produces are similar enough to those of your local collaborators that the algorithms you plan should be adaptable to them.

I suspect that phenotypes might have been overlooked in our previously collected image sets that could be found using image analysis algorithms that could better measure the physical parameters of differently sized worms. We would be eager to use algorithms your group develops to re-evaluate our past images and uncover these samples for further experimentation in our group. If this is

Room 1202, Donnelly Centre for Cellular and Biomolecular Research, 160 College St., Toronto, Ontario M5S 3E1
(Office) 416-946-8395 (Fax) 416-978-8287 (email) peter.roy@utoronto.ca

successful, of course it would allow us to rely more on automated algorithms in the future screens our group envisions. Being able to extract quantitative information from our images would also be useful for collaboration with William Ryu, a physicist recently recruited to our institution who is interested in using computational approaches to dissect worm behavior and build mathematical models. Having this new capability in our toolbox would clearly be enormously useful to our research.

Sincerely,



Peter J. Roy, Ph.D.

Room 1202, Donnelly Centre for Cellular and Biomolecular Research, 160 College St., Toronto, Ontario M5S 3E1
(Office) 416-946-8395 (Fax) 416-978-8287 (email) peter.roy@utoronto.ca

MASSACHUSETTS GENERAL HOSPITAL

HARVARD MEDICAL SCHOOL

Department of Molecular Biology
Richard B. Simches Research Center
185 Cambridge Street, 7th Floor
Boston, MA 02114
Tel: 617.726.5959, Fax: 617.726.5949
E-mail: ruvkun@molbio.mgh.harvard.edu

Gary Ruvkun, Ph.D.
Professor of Genetics
Harvard Medical School
Massachusetts General Hospital

January 25, 2010

Carolina Wahlby
Broad Institute of Harvard and MIT
Uppsala University

Dear Carolina,

I am writing to confirm our collaboration to develop image analysis software and algorithms to identify and measure *C. elegans* in high-throughput microscopy images.

As you know, my group has been conducting large-scale RNAi screens probing various pathways related to metabolism and aging, as well as the mechanisms underlying microRNA and RNAi. The power of these screens has been phenomenal, uncovering dozens of important genes with implications for human health that have been further investigated in my laboratory and the laboratories of scientists in my group.

Together with Fred Ausubel, I have worked to create an automated high-throughput *C. elegans* screening center, using fluorescence microscopy as the primary readout given the importance of visual phenotypes in the analysis of complex biological processes like metabolism and aging. The one remaining bottleneck is image analysis. While the Carpenter group successfully completed an automated imaging screen for *C. elegans* viability in collaboration with Fred's group, my group has dozens of assays that remain intractable and must be analyzed by eye.

I am therefore very eager to collaborate in your current proposal, "Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism", to create algorithms to address an important fat-regulation screen in development in my group. Because we have been collaborating with the Imaging Platform for three years and with you for the past year, assisting us on optimizing automated sample preparation and microscopy for optimal image analysis, we have been able to produce high-quality images that await your algorithm development.

As you know, all of the images we have produced are available for your work, and we are prepared to use funding from other sources to complete assay development and genome-scale RNAi screening for regulators of fat accumulation and metabolism using Nile Red staining, as well as the laboratory work to follow up on hits from this screen. We expect this screen to reveal novel regulators of fat deposition and elimination, and are hopeful that many of these genes will be useful targets for clinical obesity treatment, given the high conservation between *C. elegans* and humans.

We plan to contribute to your proposal by annotating control images for the Nile Red assay and transferring images from the full-scale screen to your group. We plan to participate interactively in the development of algorithms for this assay by helping you to assess the accuracy of automated scoring. In the course of this collaboration, my researchers who are working on the project will meet with you at least once monthly, myself included twice a year for project progress meetings.

Beyond the scope of your proposed grant to complete this particular screen, my laboratory has

dozens of other large-scale screens that will make use of the algorithms and software you develop, and we therefore are most supportive of continuing our collaboration with an emphasis now on automated image analysis for *C. elegans*.

Best wishes,

Gary Ruvkun
Professor of Genetics

High-Throughput Neurotechnology Laboratory

Massachusetts Institute of Technology

Mehmet Fatih Yanik, PhD
Assoc. Professor
Room 36-834

yanik@mit.edu
617.253.1583—Tel
617.258.5846—Fax

Carolina Wahlby, Ph.D.
Broad Institute of Harvard and MIT
7 Cambridge Center, Cambridge, MA 02142

Dear Carolina, Jan 27, 2010

I am writing to support your NIH proposal, "Image analysis algorithms for high-throughput *C. elegans* assays to identify regulators of infection and metabolism". As you know, we recently developed the first on-chip high-throughput *C. elegans* screening technologies using microfluidics. This technology allows us to immobilize awake animals within fractions of a second to take sub-cellular resolution images as well as to perform femtosecond laser surgery for genome-wide neural regeneration/degeneration studies. A great challenge however still lies in the analysis of highly rich data we acquire. Several algorithms you propose to develop for multi-well plate-based high-throughput screening could be very useful to my group's research goals over the next years.

As we have discussed last time, it will help our work considerably for your group to develop and systematize *C. elegans*-specific measurements, and to develop open-source algorithms for these in CellProfiler. We are also enthusiastic about testing in *C. elegans* the machine-learning methods that you have successfully used for complex cellular phenotypes. Several of my group members (e.g., Chris Rohde, whom you've worked with) will be able to adapt the open-source code as needed. In particular, the modular code you develop will allow us to run analysis on-the-fly by building your modules into the image acquisition pipeline. Automating the image analysis and extracting information from each sample will enable us to pursue rich phenotypic characterization of neural degeneration and regeneration in whole, living organisms. Your proposal is an excellent synergy to our technology development efforts and biological goals. We look forward to your completion of this work and its application to some of our ongoing research!

Best wishes with your grant application,

Mehmet Fatih Yanik, Ph. D.
Associate Professor
Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
Program of Computational and Systems Biology
Broad Institute of MIT and Harvard

High-Throughput Neurotechnology Laboratory

Massachusetts Institute of Technology

Mehmet Fatih Yanik, PhD
Assoc. Professor
Room 36-834

yanik@mit.edu
617.253.1583—Tel
617.258.5846—Fax

Biography: Dr. Yanik received his BS and MS at MIT Engineering in 2000, and PhD as Stanford Applied Physics Fellow in 2006. He completed short postdoctoral work in Stanford Bioengineering and Neurosurgery. He is currently serving as Assoc. Prof. at MIT's Department of Engineering. His work on high-throughput technologies, ultrafast optics, microfluidics, neuronal regeneration, coherent photonics is recognized by *NIH Director's Innovator Award*, *Packard Award in Engineering*, *Alfred Sloan Award in Neuroscience*, *NIH Eureka* (Exceptional Unconventional Research Enabling Knowledge Acceleration) *Award*, *Shillman Career Award*, *NSF Career Award*, *Silicon Valley Innovator's Challenge Award*, Technology Review Magazine's "*World's top 35 innovators under age 35*", Junior Chamber International's "*Outstanding Young Person*", and Technology Research News Magazine's "*Top ten advances of the year*". His studies have been highlighted in *ABC, The Economist, Scientific American, Nature, New Scientist, Biophotonics International, Popular Mechanics, Genome Technology*, and others. www.rle.mit.edu/Yanik

Resource Sharing Plans

It is the mission of the Broad Institute to create comprehensive tools for genomics, proteomics, chemical biology, and medicine and make them broadly available to the scientific community. In keeping with this mission, we have made plans for comprehensive sharing of both the data (images) and software produced.

Data Sharing Plan

All the raw microscopy images for the three screens will be made available to other researchers. In addition, for each assay we will compile a representative set of images and their metadata and make them available as part of the Broad Bioimage Benchmark Collection (http://www.broadinstitute.org/bbbc/) so that other researchers can download them and use them for algorithm development, validation, and comparison.

Software Sharing Plan

Software implementations of the algorithms will be incorporated into the Carpenter group's CellProfiler system, which also enables their use via ImageJ (See Research Strategy Section C, *Preliminary Studies*, and *Software development, dissemination, resource sharing, and reproducible research* for details.) In order to encourage participation by the community in its development, CellProfiler is available under the GNU General Public License (GPL) version 2 (http://www.gnu.org/licenses/old-licenses/gpl-2.0.txt). This is an open-source license with the following properties:

1. It makes the software freely available to everyone.
2. It permits the dissemination and commercialization of enhanced or customized versions of the software, or incorporation of the software or pieces of it into other software packages, with the restriction that the source code of the resulting software also be made available under the GPL. If this restriction becomes a limiting factor, we are willing to dual-license the software, i.e., grant a commercial license while continuing to release under the GPL for the benefit of the community. If incorporating CellProfiler or pieces of it into another open-source, non-GPL'ed software package would be beneficial to the scientific community, we will consider changing CellProfiler's license to the less restrictive BSD license. Note that CellProfiler's plug-in architecture allows many kinds of enhancements to be written as independently distributed modules, to which the GPL's restriction does not apply.
3. It permits anyone to continue development of the software. This preserves utility to the community in the event that we are unable or unwilling to continue development.
4. It permits anyone to modify the source code and to share modifications with others. We plan to continue to develop and release new official versions of the software, and have a track record of doing so: in the past 4 years, we have released 19 versions of CellProfiler.

We are more than happy to manage and disseminate the improvements or customizations of our tools by others by incorporating improvements made by others into the official versions of our software. To facilitate such community involvement, we have set up a publicly available source code repository and developer mailing list. We also plan to make our bug-tracking system publicly available.

CellProfiler has a plug-in architecture that makes it easy for users to enhance the software. Several labs and companies around the world have written their own CellProfiler plug-ins ("modules"), e.g., to interface with an instrument, perform a specialized analysis for their experiment, or tie into their infrastructure for data storage and analysis.

Timeline: The software will be made available as soon as it is tested and validated, and before publication. Congruent with the timeline in Section C, the software for Aim 1 should be available by the end of the second year, the software for Aim 2 by the end of the third year, and the software for Aim 3 by the end of the fifth year.

PHS 398 Checklist

OMB Number: 0925-0001

1. Application Type:

From SF 424 (R&R) Cover Page. The responses provided on the R&R cover page are repeated here for your reference, as you answer the questions that are specific to the PHS398.

* Type of Application:

☒ New ☐ Resubmission ☐ Renewal ☐ Continuation ☐ Revision

Federal Identifier:

2. Change of Investigator / Change of Institution Questions

☐ Change of principal investigator / program director

Name of former principal investigator / program director:

Prefix:
* First Name:
Middle Name:
* Last Name:
Suffix:

☐ Change of Grantee Institution

* Name of former institution:

3. Inventions and Patents (For renewal applications only)

* Inventions and Patents: Yes ☐ No ☐

If the answer is "Yes" then please answer the following:

* Previously Reported: Yes ☐ No ☐

4. * Program Income

Is program income anticipated during the periods for which the grant support is requested?

☐ Yes ☒ No

If you checked "yes" above (indicating that program income is anticipated), then use the format below to reflect the amount and source(s). Otherwise, leave this section blank.

*Budget Period	*Anticipated Amount ($)	*Source(s)

5. * Disclosure Permission Statement

If this application does not result in an award, is the Government permitted to disclose the title of your proposed project, and the name, address, telephone number and e-mail address of the official signing for the applicant organization, to organizations that may be interested in contacting you for further information (e.g., possible collaborations, investment)?

☐ Yes ☒ No