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

NAME: J. Michael Cherry

eRA COMMONS USER NAME (credential, e.g., agency login): cherry.mike

POSITION TITLE: Professor

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

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Purdue University, West Lafayette, Indiana	B.S.	1979	Biological Sciences
Purdue University, West Lafayette, Indiana	B.S.	1979	Biochemistry
University of California, Berkeley	Ph.D.	1985	Molecular Biology
Harvard University, Cambridge, Massachusetts	Postdoc	1985-1988	Genetics

A. Personal Statement

The primary goal of my research focuses on the integration of biological knowledge and published experimental results into software/database environments that provide access to these data with sophisticated tools. My work is also focused on the development of the ontologies that unify biological nomenclature by aiding curation and computation and the subsequent public distribution. I direct a group of Scientific Biocurators, Software Engineers, Data Wranglers and Bioinformatic Analysts that serve the biological research and teaching communities via our public resources. I have been active in bioinformatics since 1990 and Director of the *Saccharomyces* Genome Database (SGD) a major model organism database since 1993. My current work facilitates research by others through our collection, curation, integration and dissemination of experimental results for the budding yeast, *Saccharomyces cerevisiae*. I am currently PI for phase 3 of the NIH ENCODE project's Data Coordination Center (DCC). The ENCODE DCC is processes as much data per year as generated in all five years of the last phase of ENCODE. My accomplishments over the past year, as detailed below, have been to maintain SGD and Gene Ontology as premier resources providing manually curated "gold standard" annotations, to create and maintain the ENCODE Portal with association programmatic submission systems, and to create of expert curation and web portal interfaces for the newly Clinical Genome Project (ClinGen). My role on all the above projects is to lead an experienced group of professionals to provide the best resources possible to facilitate biomedical research. For this ClinGen proposal my role is to participate in the development of robust software that employs software and bioinformatics standards, an integrated database for the information collected by the expert curators that are associated with the clinical disease working groups for genes and variants to assign clinical relevancy to disease. That ClinGen work facilitates the operation of the many ClinGen member groups, and to create resource to aid biomedical research by providing the best quality information and interfaces possible.

B. Positions and Honors
Positions and Employment
1985 – 1988   Research Fellow in Genetics, Harvard University
1988 – 1993   Director of Computing, Molecular Biology, Massachusetts General Hospital
1991 – 1993   Research Associate in Genetics, Harvard University
1991 – 1993   Project Manager, AAtDB (An *Arabidopsis thaliana* Database). Massachusetts General Hospital & Department of Genetics, Harvard University (PI: Howard Goodman)

1993 – 1996	Head, Computing. Stanford DNA Sequence & Technology Center, NIH grant to Stanford University (PI: Ron Davis)
1993 – 2001	Chief Curator & Director, *Saccharomyces* Genome Database, NIH grant to Stanford University (PI: David Botstein)
1998 – 2000	Director, Stanford Microarray Database, Stanford University, NIH grant to Stanford University (co-PIs: Pat Brown & David Botstein)
2001 – 2013	Associate Professor (Research) of Genetics, Stanford University
2013 –	Professor (Research) of Genetics, Stanford University

Other Experience and Professional Memberships

2002 –	Member Advisory Committee, WormBase, *Caenorhabditis* database, Caltech PI: Paul Sternberg
2002 – 2006	Member, NIH Genome Research Review Committee (GRRC)
2003 – 2005	Member of Scientific Advisory Group, The Blueprint Initiative (BIND), Toronto PI: Chris Hogue
2004 – 2008	Member Advisory Committee, TIGR Rice Genome Annotation Project, PI: Robin Buell
2004 – 2008	Member Advisory Committee, EcoCyc Project at SRI. PI: Peter Karp.
2004 – 2008	Chair, Advisory Committee, Medicago Genome Sequencing, PI: Nevin Young, U. Minnesota
2005 – 2007	Scientific Advisory Panel, NHGRI ENCODE Project
2007 – 2011	External Consultants Panel, NHGRI for ENCODE and modENCODE Projects
2007 – 2009	Scientific Advisory Board, Integrative Biology Project, University of Toronto, PI: B. Andrews
2008 –	Scientific Advisory Board, FlyBase, Harvard University PI: W. Gelbart
2008 – 2012	Member, NIH Genomics, Computational Biology and Technology study section (GCAT)
2009 –	Member, Stanford Center for Genomics and Personalized Medicine (CGPM)
2010 –	Member Executive Board, International Society of Biocuration, re-elected in 2012 & 2014
2010 –	Associate Editor, Database - The Journal of Biological Databases and Curation
2011 –	Associate Editor, G3: Genes	Genomes	Genetics
2012 –	Member, Stanford Center of Computational, Evolutionary, and Human Genomics (CEHG)
2015 –	External Scientific Advisers, 4D-Nucleome Common Fund Project, NIH
2015 –	Scientific Advisory Board, FaceBase, NIH NIDCR, University of Southern California, PI: C. Kesselman
2015 –	President, International Society of Biocuration
2015 –	Scientific Advisory Board, Laboratory of Neuro Imaging Resource (LONIR), NIH, University of Southern California, PI: A. Toga
2016 –	Scientific Advisory Board, XenBase, University of Calgary, Canada, NIH NICHD, PI: P. Vize

C. Contribution to Science

My research involves identifying, validating and integrating scientific information into encyclopedic databases essential for investigation as well as scientific education. Published results of scientific experimentation are a foundation of our understanding of the natural world and provide motivation for new experiments. The combination of in-depth understanding reported in the literature with computational analyses is an essential ingredient of modern biological research. Mastery of the volumes of published literature requires comprehensive databases that provide the facts and underlying experimental data in publically accessible ways. Curation, extraction and sorting of factual experimental data, of peer-reviewed journal articles is necessary to acquire these data from its source. Large quantitative datasets using global studies extend our knowledge of genes, their products and their interactions. By integrating quantitative datasets with curated focused experimental results creates unique comprehensive databases. My group creates such essential databases and makes them available to scientists and educators seeking to understand experimental results and to teach scientific knowledge.

1. This work has been focused on the creation of the *Saccharomyces* Genome Database (www.yeastgenome.org).
    a. Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hirschman, J.E., Hitz, B.C., Karra, K., Krieger, C.J., Miyasato, S.R., Nash, R.S., Park, J., Skrzypek, M.S., Simison, M., Weng, S., Wong,

E.D. (2012) Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 40(Database issue):D700-5. doi: 10.1093/nar/gkr1029. PMCID: PMC3245034.

b. Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S., Weng, S., Wong, E.D., Lloyd, P., Skrzypek, M.S., Miyasato, S.R., Simison, M., and J.M. Cherry. (2014) The reference genome sequence of Saccharomyces cerevisiae: then and now. G3 (Bethesda). 4(3):389-98. PMCID: PMC3962479.

c. Song, G., Dickins, B.J., Demeter, J., Engel, S., Gallagher, J., Choe, K., Dunn, B., Snyder, M., and J.M. Cherry. (2015) AGAPE (Automated Genome Analysis PipelinE) for pan-genome analysis of Saccharomyces cerevisiae. PLoS One. 10(3):e0120671. PMCID: PMC4363492.

d. Engel, S.R., Weng, S., Binkley, G., Paskov, K., Song, G., and J.M. Cherry (2016) From one to many: expanding the Saccharomyces cerevisiae reference genome panel. Database (Oxford). doi: 10.1093/database/baw020. Print 2016. PMCID: PMC4795930.

e. Hellerstedt, S.T., Nash,, R.S., Weng, S., Paskov, K.M., Wong, E.D., Karra, K., Engel, S.R., Cherry, J.M. (2017) Curated protein information in the Saccharomyces Genome Database. *Database (Oxford)*. doi: 10.1093/database/bax011 PMCID: PMC5467551

f. Skrzypek, M.S., Nash, R.S., Wong, E.D., MacPherson, K.A., Hellerstedt, S.T., Engel, S.R., Karra, K., Weng, S., Sheppard, T.K., Binkley, G., Simison, M., Miyasato, S.R. and J.M. Cherry. (2018) Saccharomyces Genome Database informs human biology. Nucleic Acids Res. (Database Issue) doi: 10.1093/nar/gkx1112.

2. The development and annotation using the Gene Ontology (www.geneontology.org).
    a. Botstein, D. and J.M. Cherry. (1997) Molecular Linguistics: Extracting Information from gene and protein sequences. *Proc. Natl. Acad. Sci. USA* 94:5506-5507 PMCID: PMC34160

    b. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and G. Sherlock. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genetics* 25:25-29 PMCID: PMC3037419

    c. Balakrishnan, R., Harris, M.A., Huntley, R., Van Auken, K., and J.M. Cherry. (2013) A guide to best practices for Gene Ontology (GO) manual annotation. Database (Oxford). doi: 10.1093/database/bat054. Print 2013. PMCID: PMC3706743.

3. Tools for the analysis of gene function through the use of Gene Ontology.
    a. Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and G. Sherlock. (2004) GO::TermFinder – open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics*. 20:3710-3715 PMCID: PMC3037731

    b. Tian, W., Zhang, L.V., Tasan, M., Gibbons, F.D., King, O.D., Park, J., Wunderlich, Z., Cherry, J.M. and F.P. Roth. (2008) Combining guilt-by-association and guilt-by-profiling to predict Saccharomyces cerevisiae gene function. *Genome Biol.*, 9: Suppl 1:S7 PMCID: PMC2447541

    c. Park, J., Costanzo, M.C., Balakrishnan, R., Cherry, J.M. and E.L. Hong. (2012) CvManGO, a method for leveraging computational predictions to improve literature-based Gene Ontology annotations. *Database (Oxford)*. doi:10.1093/database/bas001 Print 2012. PMCID: PMC3308158

    d. Dutkowski, J., Kramer, M., Surma, M.A., Balakrishnan, R., Cherry, J.M., Krogan, N.J., and Ideker, T. (2012) A gene ontology inferred from molecular networks. *Nat. Biotechnol.* 31:38-45. PMCID: PMC3654867.

4. Integration of online resources and gene product annotations
    a. Rangarajan, A., Schedl, T., Yook, K., Chan, J., Haenel, S., Otis, L., Faelten, S., Depellegrin-Connelly, T., Isaacson, R., Skrzypek, M.S., Marygold, S.J., Stefancsik, R., Cherry, J.M., Sternberg, P.W. and H.M. Müller (2011) Toward an interactive article: integrating journals and biological databases. BMC Bioinformatics. 2011: 12 175 PMCID: PMC3213741

    b. Klionsky, D.J., Bruford, E.A., Cherry, J.M., Hodgkin, J., Laulederkind, S.J. and Singer, A.G. (2012) In the beginning there was babble… *Autophagy*. 8:1165-1167 PMCID: PMC3625114

    c.   Boyle, A., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., Cherry, J.M. and M. Snyder (2012) Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research*. 22:1790-1797 PMCID: PMC3431494

    d.   Sullivan, J., Karra, K., Moxon, S.A., Vallejos, A., Motenko, H., Wong, J.D., Aleksic, J., Balakrishnan, R., Binkley, G., Harris, T., Hitz, B., Jayaraman, P., Lyne, R., Neuhauser, S., Pich, C., Smith, R.N., Trinh, Q., Cherry, J.M., Richardson, J., Stein, L., Twigger, S., Westerfield, M., Worthey, E., Micklem, G. (2013) InterMOD: integrated data and tools for the unification of model organism research. *Sci Rep.* doi: 10.1038/srep01802. PMCID: PMC3647165

5. Application of ontologies for enhanced annotation of metadata for experimental results, and the use of metadata to facilitate enhanced access to these data.

    a.   Malladi, V.S., Erickson, D.T., Podduturi, N.R., Rowe, L.D., Chan, E.T., Davidson, J.M., Hitz, B.C., Ho, M., Lee, B.T., Miyasato, S., Roe, G.R., Simison, M., Sloan, C.A., Strattan, J.S., Tanaka, F., Kent, W.J., Cherry, J.M., and E.L. Hong. (2015) Ontology application and use at the ENCODE DCC. Database (Oxford). doi: 10.1093/database/bav010. Print 2015. PMCID: PMC4360730.

    b.   Sloan, C.A., Chan, E.T., Davidson, J.M., Malladi, V.S., Strattan, J.S., Hitz, B.C., Gabdank, I., Narayanan, A.K., Ho, M., Lee, B.T., Rowe, L.D., Dreszer, T.R., Roe, G., Podduturi, N.R., Tanaka, F., Hong, E.L., and J.M. Cherry. (2016) ENCODE data at the ENCODE portal. Nucleic Acids Res. 44(D1):D726-32. doi: 10.1093/nar/gkv1160. PMCID: PMC4702836.

    c.   Hong, E.L., Sloan, C.A., Chan, E.T., Davidson, J.M., Malladi, V.S., Strattan, J.S., Hitz, B.C., Gabdank, I., Narayanan, A.K., Ho, M., Lee, B.T., Rowe, L.D., Dreszer, T.R., Roe, G.R., Podduturi, N.R., Tanaka, F., Hilton, J.A., and J.M. Cherry. (2016) Principles of metadata organization at the ENCODE Data Coordination Center. Database (Oxford). doi: 10.1093/database/baw001.

    d.   Hitz, B., Rowe, L.D., Podduturi, N., Glick, D., Baymuradov, U., Malladi, V., Chan, E., Davidson, J., Gabdank, I., Narayanan, A., Onate, K., Ho, M., Lee, B., Miyasato, S., Dreszer, T., Sloan, C., Strattan, J.S., Tanaka, F., Hong, E.L., and J. M. Cherry. (2016) SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata. *PLoS One* doi: 10.1371/journal.pone.0175310

    e.   Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., Onate, K.C., Graham, K., Miyasato, S.R., Dreszer, T.R., Strattan, J.S., Jolanki, O., Tanaka, F.Y. and J.M. Cherry. (2018) The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. (Database Issue) doi: 10.1093/nar/gkx1081

Complete List of Published Work in MyBibliography:
http://www.ncbi.nlm.nih.gov/sites/myncbi/joe%20michael.cherry.1/bibliography/40619086/public/?sort=date&direction=descending