

JUSTIN THOMAS REESE

717 Potter St
Lawrence Berkeley National Laboratory
Berkeley, CA 94710
510-486-5078 (Office) 415-909-9418 (Mobile)
Email: justinreese@lbl.gov

SUMMARY

I am an expert in computational biology, bioinformatics, machine learning, and data science, with experience in programming, teaching, and managing scientific research projects. I have worked for sixteen years developing software, performing bioinformatic analyses, teaching, and conducting original research.

EXPERIENCE

2019 – Present

Lawrence Berkeley National Laboratory

Computer Research Scientist

- Project management/software team leadership
- Software development (Python, Java, Javascript, Rust)
- Investigated funding opportunities, grant writing
- Conducted original research on graph machine learning, structural variants, and COVID-19 drug repurposing
- Presented research and conducted outreach to encourage adoption of software developed by our team
- KG-COVID-19 project:
 - Led successful application for funding
 - Led team that designed and developed software to produce COVID-19 knowledge graph
 - Developed two novel graph machine learning techniques for drug repurposing
 - Published research in peer-reviewed journal
 - Led outreach to encourage adoption of KG into other projects (N3C, NVBL, COV-IRT)
- NVBL Area 1 project:
 - Helped organize system for coordination of protein target information across national laboratories
 - Facilitated adoption/reuse of KG-COVID-19 knowledge graph and machine learning tools to support NVBL goals of target identification and drug repurposing
- Monarch Initiative project:
 - Contributed to software development of Dipper ETL pipeline (Python) for data ingest

- Contributed to software development of Monarch user interface (Javascript)
- Contributed to software development (Java) for LIRICAL project and publication
- IDG project:
 - Project management and software development (Python) for Embiggen/NEAT ML framework
 - Contributed domain expertise in biology and machine learning to help identify synthetic lethal relationships for cancer treatment
 - Organized regular meetings and ad hoc hackathons to facilitate cross-site software development between team members at LBNL, the University of Milan, and Jackson Lab

2015 – 2019

Khon Kaen University/Vital Scientific, LLC

Adjunct Faculty/Consultant –Khon Kaen University, Citrus Research Board, The Asia Foundation, others

- Designed and developed online web portal www.datoc.org for managing data and encouraging collaboration amongst scientists and citrus growers to combat Citrus greening disease (huánglóngbìng)
- Provided technical advice, service and guidance related to computational biology, bioinformatics, statistics, and general IT to various academic researchers and industry clients
- Developed computational biology projects to aid in treatment of cholangiocarcinoma and nasopharyngeal carcinoma in coordination with Khon Kaen University faculty
- Advised faculty and students dealing with computational biology data related to existing projects, and assisted faculty and students in preparation of scientific manuscripts and presentations

2010 – 2015

Genformatic, LLC

Chief Operating Officer

- Produced three publications describing a novel algorithm for more effective genome variant detection, novel research on psyllid genome to combat Citrus greening disease (huánglóngbìng), and a practical method for genomic analysis of archived newborn bloodspot data
- Managed bioinformatics team that performed genome variant detection and bespoke bioinformatics analyses for Genformatic, LLC, generating revenue
- Led development of several bioinformatic products, including novel software for improving genome variant detection, custom SNP and indel calling pipeline and genome variant annotation pipeline
- Assisted in business development

2012 – 2013

University of Missouri

Assistant Research Professor

- Lead developer of bovinegenome.org and hymenopteragenome.org, web portals to display information pertaining to the *Bos taurus* and hymenopteran (e.g. *Apis mellifera*, *Nasonia vitripennis*) genomes
- Developed novel software and designed/maintained IT infrastructure for the annotation of genes in three genomes: *Apis mellifera*, *Nasonia vitripennis*, *Bos taurus*
- Developed scientific software for research and analysis of several newly sequenced genomes, including *Apis mellifera*, *Bos taurus*, *Strongylocentrotus purpuratus*, *Nasonia vitripennis*, *Tribolium castaneum*, *Atta cephalotes*, *Pogonomyrmex barbatus*, *Linepithema humile* and *Pediculus humanus*.

- Lectured on bioinformatics and scientific programming

2006 – 2012

Georgetown University

Affiliate Research Professor

- Lead developer of bovinegenome.org and hymenopteragenome.org web portals (see above)
- Developed novel software and designed/maintained IT infrastructure for the annotation of genes in three genomes: *Apis mellifera*, *Nasonia vitripennis*, *Bos taurus*
- Developed scientific software for research and analysis of several newly sequenced genomes, including *Apis mellifera* and others (see above)
- Lectured on bioinformatics and scientific programming

2004 – 2006

Texas A&M University

Research Associate

- Lead developer of hymenopteragenome.org (see above)
- Training of graduate students and postdoctoral staff as required

2004 – 2005

First Global Community College, Nong Khai, Thailand

Lecturer/Director of Science (Part-time Volunteer)

- Set up and administered a science and IT program consisting of six teachers
- Taught science and computer classes

1998 – 2004

University of Virginia (Laboratory of William Pearson)

Graduate student

- Investigated the evolution of repeated sequences in protein and DNA
- Developed a mathematical model to explain the evolution of proteins in bacterial genomes
- Developed formula for optimal gap penalties for sequence alignment and database searching, improving sensitivity and selectivity of sequence analysis tools (FASTA, BLAST)

EDUCATION

Ph.D.	University of Virginia	May 2004
Expertise: Computational biology and bioinformatics		
M.S.	University of Georgia	August 1997
Expertise: Immunology		
B.S.	Clemson University	December 1994
Major: Biochemistry (<i>cum laude</i>)		

SKILLS

- Experienced in scientific project and software project management
- Experienced in teaching and mentoring at the undergraduate and graduate level
- Experienced in good programming practice
- Experienced in machine learning, especially graph machine learning
- Proficient in several programming languages (Python, Java, Ruby, R, C, Rust)
- Proficient in web development and database programming

LANGUAGES

- English: native speaker, excellent speaking and writing skills
- Thai: fluent in spoken Thai, basic reading and writing skills
- Lao: working knowledge of spoken and written Lao
- Spanish: rudimentary knowledge of spoken and written Spanish

PUBLICATIONS

Reese JT, Coleman B, Chan L, Blau H, Callahan TJ, Cappelletti L, Fontana T, Bradwell KR, Harris NL, Casiraghi E, Valentini G, Karlebach G, Deer R, McMurry JA, Haendel MA, Chute CG, Pfaff E, Moffitt R, Spratt H, Singh JA, Mungall CJ, Williams AE, Robinson PN. [NSAID use and clinical outcomes in COVID-19 patients: a 38-center retrospective cohort study](#). *Virol J*. 2022 May 15;19(1):84. doi: 10.1186/s12985-022-01813-2. PubMed PMID: 35570298.

Coleman, B., Casiraghi, E., Blau, H., Chan, L., Haendel, M. A., Laraway, B., Callahan, T.J., Deer, R., Wilkins, K.J., **Reese, J.**, Robinson, P. (2022) Risk of new-onset psychiatric sequelae of COVID-19 in the early and late post-acute phase. *World Psychiatry* 07 May. <https://doi.org/10.1002/wps.20992>

Jacobsen J., Kelly, C., Cipriani, V., Genomics England Research Consortium, Mungall, C., **Reese, J.**, Danis, D., Robinson, P., Smedley, D. (2022) Phenotype-driven approaches to enhance variant prioritization and diagnosis of rare disease. *Hum Mutat*. Apr 7. doi:10.1002/humu.24380

Cappelletti, L., Fontana, T., Casiraghi, E., Ravanmehr, V., Callahan, T., Joachimiak, M., Mungall, C., Robinson, P., **Reese, J.**, Valentini, G. (2021) GraPE: fast and scalable Graph Processing and Embedding arXiv. 2021. <https://arxiv.org/abs/2110.06196>

Danis D, Jacobsen JOB, Balachandran P, Zhu Q, Yilmaz F, **Reese J**, Haimel M, Lyon GJ, Helbig I, Mungall CJ, Beck C, Lee C, Smedley D, Robinson PN. (2022). SvAnna: efficient and accurate pathogenicity prediction of coding and regulatory structural variants in long-read genome sequencing. *Genome Med* . Apr 28;14(1):44. doi: 10.1186/s13073-022-01046-6

Ravanmehr V, Blau H, Cappelletti L, Fontana T, Carmody L, Coleman B, George J, **Reese J**, Joachimiak M, Bocci G, Hansen P, Bult C, Rueter J, Casiraghi E, Valentini G, Mungall C, Oprea TI, Robinson PN. [Supervised learning with word embeddings derived from PubMed captures latent knowledge about protein kinases and cancer](#). *NAR Genom Bioinform*. 2021 Dec;3(4):lqab113. doi: 10.1093/nargab/lqab113. eCollection 2021 Dec. PubMed PMID: 34888523; PubMed Central PMCID: PMC8652379.

Rando, H. M., Bennett, T. D., Byrd, J. B., Bramante, C., Callahan, T. J., Chute, C. G., Davis, H. E., Deer, R., Gagnier, J., Korashy, F. M., Liu, F., McMurry, J.A., Moffitt, R.A., Pfaff, E.R., **Reese, J.T.**, Relevo, R., Robinson, P.N., Saltz, J.H., Solomonides, A., Sule, A., Topaloglu, U., Haendel, M. A. (2021). Challenges in defining Long COVID: Striking differences across literature, Electronic Health Records, and patient-reported information. *medRxiv*. doi: <https://doi.org/10.1101/2021.03.20.21253896>

Valentini G, Casiraghi E, Cappelletti L, Ravanmehr V, Fontana T, **Reese J**, Robinson P. Het-node2vec: second order random walk sampling for heterogeneous multigraphs embedding. arXiv [cs.LG]. 2021. <http://arxiv.org/abs/2101.01425>

Deer RR, Rock MA, Vasilevsky N, Carmody LC, Rando HM, Anzalone AJ, Callahan TJ, Bramante CT, Chute CG, Greene CS, Others, **Reese JT**, Others. Characterizing Long COVID: Deep Phenotype of a Complex Condition. *medRxiv*. Cold Spring Harbor Laboratory Press; 2021

Reese, J.T. as part of N3C Consortium. (2021) The National COVID Cohort Collaborative (N3C): Rationale, design, infrastructure, and deployment. *JAMIA*, V28:3, 427–443.

Reese, J.T., Unni, D.R., Callahan, T.J., Cappelletti, L., Ravanmehr, V., Carbon, S., Shefchek, K.A., Good, B.M., Balhoff, J.P., Fontana, T., Blau, H., Matentzoglou, N., Harris, N.L. and Munoz-Torres, M.C., Haendel, M.A., Robinson, P.N., Joachimiak, M.P., and Mungall, C.J. (2020). KG-COVID-19: a framework to produce customized knowledge graphs for COVID-19 response. *Patterns*, 100155.

Shefchek KA, Harris NL, Gargano M, Matentzoglou N, Unni D, Brush M, Keith D, Conlin T, Vasilevsky N, Zhang XA, Balhoff JP., **Reese, J.T.**, et al. (2020) The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nucleic acids research*. (D1) 704-15.

Casiraghi, E., Malchiodi, D., Trucco, G., Frasca, M., Cappelletti, L., Fontana, T., Esposito, A.A., Avola, E., Jachetti, A., **Reese, J.** and Rizzi, A., (2020). Explainable machine learning for early assessment of COVID-19 risk prediction in emergency departments. *IEEE Access*.

Robinson, P.N., Ravanmehr, V., Jacobsen, J.O., Danis, D., Zhang, X.A., Carmody, L., Gargano, M., Thaxton, C., **Reese, J.**, Holtgrewe, M. and Koehler, S., 2020. Interpretable Clinical Genomics with a Likelihood Ratio Paradigm. *Am. J. Hum. Genet.* 107(3), 403-417.

Wuttiumporn, K., Pasurivong, O., Boonsawat, W., **Reese, J. T.**, Intarapoka, B., & Khisanapant, W. (2018). Associations between Airway Inflammation and Indices of Sleep Apnea Severity in Obstructive Sleep Apnea Patients. *Srinagarind Medical Journal*, 33(2), 102-109.

Tarver, M., Huang, Q., de Guzman, L., Rinderer, T., Holloway, B., **Reese, J.**, Weaver, D., and Evans, J. (2016). Transcriptomic and functional resources for the small hive beetle *Aethina tumida*, a worldwide parasite of honey bees. *Genom Data*. Sep; 9: 97–99.

Cantarel, B., Lei, Y., Weaver, D., Zhu, H., Farrell, A., Benstead-Hume, G., **Reese, J.** and Finnell, R. (2015). Analysis of archived residual newborn screening blood spots after whole genome amplification. *BMC Genomics*, 16:602.

Cantarel, B., Weaver, D., McNeill, N., Zhang, Jianhua, Z., Mackey, A., **Reese, J.** (2014). BAYSIC: a Bayesian method for combining sets of genome variants with improved specificity and sensitivity. *BMC Genomics*, 15(1).

Terrapon, N., Li, C., Robertson, H. M., Ji, L., Meng, X., Booth, W., Chen, Z., Childers, C., Glastad, K., Gokhale, K., Gowin, J., Gronenberg, W., Hermansen, R., Hu, H., Hunt, B., Huylmans, A., Khalil, S., Mitchell, R., Munoz-Torres, M., Mustard, J., Pan, H., **Reese, J.**, Scharf, M., Sun, F., Vogel, H., Xiao, J., Yang, W., Yang, Z., Yang, Z., Zhou, J., Zhu, J., Brent, C., Elsik, C., Goodisman, M., Liberles, D., Roe, R. M., Vargo, E., Vilcinskis, A., Wang, J., Bornberg-Bauer, E., Korb, J., Zhang, G. and Liebig, J. (2014). Molecular traces of alternative social organization in a termite genome. *Nature communications*, 5.

Elsik, C., Worley, K., Bennett, A., Beye, M., Camara, F., Childers, C., de Graff, Debyser, G., Deng, J., Devreese, B., Elhaik, E., Evans, J., Foster, L., Graur, D., Guigo, R., Hoff, K., Holder, M., Hudson, M., Hunt, G., Jiang, H., Joshi, V., Khetani, R., Kosarev, P., Kovar, C., Ma, J., Maleskzka, R., Moritz, R., Muñoz-Torres, M., Murphy, T., Muzny, M., Newsham, F., **Reese, J.**, Robertson, M., Robinson, G., Rueppel, O., Solovyev, V., Stanke, M., Stolle, E., Tsuruda, J., Van Vaerenbergh, M., Waterhouse, R., Weaver, D., Whitfield, C., Wu, Y., Zdobnov, E., Lan, Z., Zhu, D. and Gibbs, R. (2014). Finding the missing honey bee genes: lessons learned from a genome upgrade. *BMC Genomics*, 15(1), 86.

Reese, J., Christenson, M., Leng, N., Saha, S., Cantarel, B., Lindeberg, M., Tamborindeguy, C., MacCarthy, J., Weaver, D., Trease, A., Ready, S., Davis, V., McCormick, C., Haudenschild, C., Han, S., Johnson, S., Shelby, K.,

Huang, H., Bextine, B., Shatters, R., Hall, D., Davis, P. and Hunter, W. (2013). Characterization of the Asian Citrus Psyllid Transcriptome. *J Genomics*, 2:54-58.

Lee, E., Helt, G., **Reese, J.**, Muñoz-Torres, M., Childers, C., Buels, R., Stein, L., Holmes, I., Elsik, C. and Lewis, S. (2013). Web Apollo: A Web-based Genomic Annotation Editing Platform. *Genome Biology*, 14, R93.

Simola, D. F., Wissler, L., Donahue, G., Waterhouse, R.M., Helmkampf, M., Roux, J., Nygaard, S., Glastad, K., Hagen, D., Viljakainen, L., **Reese, J.**, Hunt, B., Graur, D., Elhaik, E., Kriventseva, E., Wen, J., Parker, B., Cash, E., Privman, E., Childers, C., Muñoz-Torres, M., Boomsma, J., Bornberg-Bauer, E., Currie, C., Elsik, C., Suen, G., Goodisman, M., Keller, L., Liebig, J., Rawls, A., Reinberg, D., Smith, C.D., Smith, C.R., Tsutsui, N., Wurm, Y., Zdobnov, E., Berger, S. and Gadau, J. (2013). Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Research*, 23(8), 1235-1247.

Saha, S., Hunter, W. B., **Reese, J.**, Morgan, J. K., Marutani-Hert, M., Huang, H., and Lindeberg, M. (2012). Survey of endosymbionts in the Diaphorina citri metagenome and assembly of a Wolbachia wDi draft genome. *PLoS One*, 7(11), e50067.

Smith, C.D., Abouheif, E., Benton, R., Croset, V., Currie, C.R., Elhaik, E., Elsik, C.G., Fave, M-J, Fernandes, V., Gadau, J., Gibson, J.D., Graur, D., Hagen, D.E., Helmkampf, M., Holt, C., Hu, H., Johnson, B.R., Johnson, R.M., Abderrahman, K., Kim, J.K., Mathis, K.A., Muñoz-Torres, M.C., Murphy, M.C., Nakamura, R., Nigam, S., Overson, R., Placek, J., Rajakumar, R., **Reese, J.T.**, Robertson, H.M., Smith, C.R., Suen, G., Tao, S., Torres, C.W., van Wilgenburg, E., Viljakainen, L., Vinniegra, A-S. I., Walden, K.K.O., Yandell, M.D., Zimin, A. & N.D. Tsutsui (2011). The Genome of the Globally Widespread and Invasive Argentine Ant (*Linepithema humile*). *Proceedings of the National Academy of Sciences USA*. 108:5667-5672.

Smith, C.R., Smith, C.D., Robertson, H.M., Helmkampf, M., Zimin, A., Yandell, M., Holt, C., Hu, H., Abouheif, E., Benton, R., Cash, E., Croset, V., Currie, C.R., Elhaik, E., Elsik, C.G., Fave, M-J., Fernandes, V., Gibson, J.D., Graur, D., Gronenber, W., Grubbs, K.J., Hagen, D., Vinniegra, A-S.I., Johnson, B.R., Johnson, R., Khila, A., Kim, J.W., Mathis, K.A., Muñoz-Torres, M.C., Murphy, M.C., Mustard, J.A., Nakamura, R., Neihuis, O., Nigam, S., Overson, R., Placek, J., Rajakumar, R., **Reese, J.**, Suen, G., Shu, T., Torres, C.W., Tsutsui, N.D., Viljakainen, L., Wolschin, F. & J. Gadau (2011). A Draft Genome of the Red Harvester Ant, *Pogonomyrmex barbatus*: a Model for Reproductive Division of Labor and Social Complexity. *Proceedings of the National Academy of Sciences USA*.

Childers, C.,* **Reese, J.**,* Sundaram, J.,* Vile, D., Dickens, C., Childs, K., Salih, H., Bennett, A., Hagen, D., Adelson, D. and Elsik, C. (2011) Bovine Genome Database: integrated tools for genome annotation and discovery. *Nucleic Acids Res.* 1:39:D830-4.

Munoz-Torres, M.,* **Reese, J.**,* Childers, C.,* Bennett, A., Sundaram, J., Childs, K., Anzola, J., Milshina, N. and Elsik, C. (2010) Hymenoptera Genome Database: integrated community resources for insect species of the order Hymenoptera. *Nucleic Acids Res.* 39:D658-662.

Reese, J.,* Childers, C.,* Sundaram, J.,* Dickens, M., Childs, K., Vile, D. and Elsik, C. (2010) Bovine Genome Database: supporting community annotation and analysis of the *Bos taurus* genome. *BMC Genomics*. 11:645.

Reese, J., as member of The *Nasonia* Genome Working Group (2010) Functional and Evolutionary Insights from the Genomes of Three Parasitoid *Nasonia* Species. *Science*. 327:343 – 348.

Reese, J., as member of Tribolium Genome Sequencing Consortium (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature*. 2008 452:949-55

Weaver, D., Anzola, J., Evans, J., Reid, J., **Reese, J.**, Childs, K., Zdobnov, E., Samanta, M., Miller, J. and Elsik, C. (2007) Computational and transcriptional evidence for microRNAs in the honey bee genome. *Genome Biology*. 8:R97.

Robertson, H., **Reese, J.**, Milshina, N., Agarwala, R., Solignac, M., Walden, K., and Elsik, E. (2007) Manual superscaffolding of honey bee (*Apis mellifera*) chromosomes 12-16: implications for the draft genome assembly version 4, gene annotation, and chromosome structure. *Insect Molecular Biology*. 16:401-410.

Elsik, E., Mackey, A., **Reese, J.**, Milshina, N., Roos, D. and Weinstock, G. (2007) Creating a honey bee consensus gene list. *Genome Biology*. 8:R13.

Reese, J., as member of Sea Urchin Genome Sequencing Consortium (2006) The Genome of the Sea Urchin *Strongylocentrotus purpuratus*. *Science*. 314:941 – 952.

Elsik, C., Worley, K., Zhang, L., Milshina, N., Jiang, H., **Reese, J.**, Childs, K., Venkatraman, A., Dickens, C., Weinstock, G., and Gibbs, R. (2006) Community annotation: Procedures, protocols, and supporting tools. *Genome Research*. 16:1329-1333.

Reese, J. as member of the Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*. 443:931-949.

Miyake, T., **Reese, J.**, Loch, C., Auble D. and Li, R. (2004) Genome-wide Analysis of ARS Binding Factor 1 (Abf1p)-Mediated Transcriptional Regulation in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 2004 Jun 2

VanLoock, M., Yu, X., Yang, S., **Reese, J.** and Egelman, E. (2004) What is the Structure of the RecA-DNA Filament? *Curr. Protein Pept. Sci.* 5:73-9

Reese, J. and Pearson, W. (2002) Empirical Determination of Effective Gap Penalties for Sequence Alignment. *Bioinformatics*. 18: 1500-7.

Reese, J., Mehta, H., Chappell, C. and Bamezai, A. (2001) Downregulated Expression of Ly-6-ThB on Developing T cells Marks CD4+CD8+ Subset Undergoing Selection in the Thymus. *Dev. Immunol.* 8(2):107-21.

PATENTS

Device, system and method for securing and comparing genomic data. US Patent US9449191B2.