

CHARLES A. TILFORD, Ph.D.

Tilford@biocode.fastmail.fm

978-341-8270

www.linkedin.com/in/charles-tilford/

github.com/maptracker (Professional)

github.com/VCF (Personal)

Skills & Expertise

- Over 18 years of experience creating, deploying and maintaining tools to store, analyze and visualize genome-scale information in biology and chemistry. Software developed from earliest R&D to clinical trials. Successful collaborations with Genomics, Target Teams, Chemistry, Legal:
 - Core designer and contributor to BMS Reproducible Research framework
 - In-house expert on R package development and management
 - Novel short-algorithm alignment tool to support ASO development
 - Polymorphism / mutation integration and analysis supporting several NGS projects
 - Detection, analysis and data warehousing of viral polymorphisms in HCV trials
 - Sequence processing pipeline for Adnexus PROfusion™ library screens
 - Chem-Bio data integration supporting design and analysis of chemogenomics screens
- Fluency in typical bioinformatics arenas: alignment algorithms; public data sources at NCBI, EBI, SIB, OBO, etc; BioPerl; commercial data sources for patents, targets, compounds.
- Skilled in R, bash, Perl, JavaScript, DHTML/AJAX/JSON. Develop, optimize and operate terabyte+ databases in PostgreSQL, MariaDB and Oracle. Create R packages and manage internal R repository. Linux is primary environment for *both* work & home. Functional in git, Shiny, AWS.
- Formal molecular biology training allows rapid understanding of colleagues' needs, resulting in advance recognition of potential problem points and faster project completion.

Professional Experience & Education

Bristol-Myers Squibb Company, Hopewell, NJ & Cambridge, MA *2001 - Present*

- Principal Scientist, TBio / Translational Medicine / Computational Genomics
- Development of software tools to aid in the characterization of drug targets. Flagship systems:
 - **myRepository** ([GitHub](#)), an R package designed to aid formal R package development, and to configure and maintain an in-house R package repository. Part of our RR initiative ([GitHub](#))
 - **MapTracker** ([GitHub](#)), a massive, generic graph database. 1.2B nodes, 2.0B edges, 3.5B metadata assignments. A PostgreSQL database weighing over a terabyte, supported by command line and web front ends, plus a modular API, in tens of thousands of lines of ooPerl. Relates biological entities, compounds, ontologies, synonyms and much more. Calculates coordinate transformations across sequence alignments. Captures both public and internal data.
 - **GenAcc** ([GitHub](#)), a Perl API allowing transitive graph queries to MapTracker. Utilizes recursive logic to discover indirect (and appropriate) linkage between distant data. Algorithmic output cached in a simple database to allow rapid recovery of precomputed results on re-query.
 - **AnnotatedMatrix** ([GitHub](#)), an R package to allow fully self-contained From/To mapping matrices with rich metadata and pre-defined filtering criteria. Rolling out as a Reproducible Research-friendly alternative to MapTracker.
 - **SetFisher** ([GitHub](#)), a framework utilizing AnnotatedMatrix to allow hypergeometric enrichment analysis that accounts for non-unique cardinality, for example Affymetrix arrays with many-to-many relationships between probesets and loci.
 - **Simple Oligo Search** ([GitHub](#)), allowing exhaustive matches for very short oligos against a genome while accommodating both common polymorphisms and mismatches that would

- prevent detection in other algorithms. Used for off-target detection in antisense oligo projects.
- **MapLoc** ([GitHub](#)), a repository for internal and external genomic polymorphism and mutation data. Aggressively normalizes variant information from multiple sources to allow integrated search and display. Allows queries against genes, genomic ranges, populations, data sources, MAF and impact. Full Perl API, reports in [CanvasXpress](#), Excel workbooks, HTML or text. Supported next-generation sequencing projects in Oncology.
 - **ChemBio Hopper** ([GitHub](#)), an aggregator for chemistry-target assay information. Allows filtered and annotated reports of potent compounds given genes, and target genes given compounds. Used in the design of a chemogenomics screening deck, and for interpreting the resulting screens. Resulted in an Applied Biotechnology Innovation Award in 2010.
 - **V-SAW** ([Overview](#)), a tool for finding polymorphisms in pooled viral sequencing reads. Filters likely mutations from sequencing noise, facilitates rapid researcher validation of candidates, and prepares high-level views that summarize altered sites, while allowing drill-down to trace. Used to track emergent polymorphisms in HCV clinical trials. Two major internal awards.
 - **Library Helper**, a sequence analysis system for [PROfusion](#)TM. Prior to acquisition by BMS, Adnexus researchers would spend many hours manually analyzing AdnectinTM sequences from each library screen. Reduces human investment to a couple minutes, provides higher fidelity analysis, and presents results in standardized excel workbooks and interactive web pages.
 - **Gene Set Enrichment** ([GitHub](#)), used to identify classifiers (Gene Ontology, MSigDB, CDD, etc) that are unusually enriched or depleted in a list of genes. Implemented in Perl, with hypergeometric distribution optimized in Inline C++. Extensively used by researchers to aid in interpretation of transcriptional profiling experiments.

Massachusetts Institute of Technology, Cambridge, MA

1994 - 2001

- Ph.D. from Biology Department, advisor Dr. David Page
- Thesis: [Mapping the Human Y Chromosome](#).
- Extensive bench work: DNA purification, cloning, PCR, library construction and screening, subtractive hybridization, agarose and polyacrylamide gel electrophoresis, blotting, capillary sequencing. BL2 cell culture with murine cell lines.
- Programming for a 20 x 96 well PCR work flow on a Packard MultiProbeTM liquid handling robot. Designed high-throughput, low cost agarose gel system for analysis of PCR products.
- Wrote Java applet to partially automate analysis of EtBr-stained gel images for the [TNG radiation hybrid panel](#). Applet identified wells with product, allowed corrections, and stored results in flat file. Perl scripts perform simulated annealing analysis of results to determine genomic arrangement of markers. Developed novel graphical methods for displaying results.
- Teaching Assistant for General Biochemistry and Introduction to Experimental Biology. Served on Whitehead Education Committee.

United States Peace Corps

1991 - 1993

- High School Science Teacher, [Visitation Boarding School](#), Makumbi, Zimbabwe
- Taught basic science to 120 O-level students and chemistry to 25 A-level students in a rural school 50 km north of capital. Organized a Science Club, Flying Models Club, and established a space library with materials donated by NASA. "Coached" basketball.
- Partially replaced corporal punishment with a system of awards that granted public recognition of students and demerits that required public works around the school grounds.

Williams College, Williamstown, MA

1987 - 1991

- B.A. in Chemistry. GPA 3.76. *Magna Cum Laude*, Highest Honors, Phi Beta Kappa, Sigma Xi
- Thesis: [Laser Spectroscopy of CF₃I Monomers and Computational Analysis of CF₃I Van Der Waals' Complexes](#), advisor Dr. Jay Thoman

Publications

- Wu Y, Broadaway KA, Raulerson CK, Scott LJ, Pan C, Ko A, He A, Tilford C, Fuchsberger C, Locke AE, Stringham HM, Jackson AU, Narisu N, Kuusisto J, Pajukanta P, Collins FS, Boehnke M, Laakso M, Lusis AJ, Civelek M, Mohlke KL. **Colocalization of GWAS and eQTL signals at loci with multiple signals identifies additional candidate genes for body fat distribution.** [Hum Mol Genet](#). 2019 Dec 15;28(24):4161-4172.
- Wang X, Tilford C, Neuhaus I, Mintier G, Guo Q, Feder JN, Kirov S. **CRISPR-DAV: CRISPR NGS Data Analysis and Visualization Pipeline.** [Bioinformatics](#). 2017 Aug 14.
- Civelek M, Wu Y, Pan C, Raulerson CK, Ko A, He A, Tilford C, Saleem NK, Stančáková A, Scott LJ, Fuchsberger C, Stringham HM, Jackson AU, Narisu N, Chines PS, Small KS, Kuusisto J, Parks BW, Pajukanta P, Kirchgessner T, Collins FS, Gargalovic PS, Boehnke M, Laakso M, Mohlke KL, Lusis AJ. **Genetic Regulation of Adipose Gene Expression and Cardio-Metabolic Traits.** [Am J Hum Genet](#). 2017 Mar 2;100(3):428-443.
- Tilford CA, Neuhaus I. **MapLoc genetic variation database.** <https://github.com/VCF/MapLoc/>
- Quan Y, Jin Y, Faria TN, Tilford CA, He A, Wall DA, Smith RL, Vig BS. **Expression Profile of Drug and Nutrient Absorption Related Genes in Madin-Darby Canine Kidney (MDCK) Cells Grown under Differentiation Conditions.** [Pharmaceutics](#). 2012 Jun 18;4(2):314-33.
- Ross-Macdonald P, de Silva H, Patel V, Truong A, He A, Neuhaus I, Tilford C, Ji R, Siemers N, Greer A, Carboni J, Gottardis M, Menard K, Lee F, Dodier M, Frennesson D, Sampognaro A, Saulnier M, Trainor G, Vyas D, Zimmermann K, Wittman M. **Biochemical and transcriptional profiling to triage additional activities in a series of IGF-1R/IR inhibitors.** [Bioorg Med Chem](#). 2012 Mar 15;20(6):1961-72.
- Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Weitz K, Camp DG 2nd, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, Gargalovic P, Eskin E, Kirchgessner T, Smith DJ, Smith RD, Lusis AJ. **Comparative analysis of proteome and transcriptome variation in mouse.** [PLoS Genet](#). 2011 Jun;7(6):e1001393.
- Romanoski CE, Lee S, Kim MJ, Ingram-Drake L, Plaisier CL, Yordanova R, Tilford C, Guan B, He A, Gargalovic PS, Kirchgessner TG, Berliner JA, Lusis AJ. **Systems genetics analysis of gene-by-environment interactions in human cells.** [Am J Hum Genet](#). 2010 Mar 12;86(3):399-410.
- Tilford C, Siemers N. **Gene Set Enrichment Analysis.** [Methods Mol Biol](#). 2009;563:99-121.
- Ross-Macdonald P, de Silva H, Guo Q, Xiao H, Hung CY, Penhallow B, Markwalder J, He L, Attar RM, Lin TA, Seitz S, Tilford C, Wardwell-Swanson J, Jackson D. **Identification of a nonkinase target mediating cytotoxicity of novel kinase inhibitors.** [Mol Cancer Ther](#). 2008 Nov;7(11):3490-8.
- Ucar D, Neuhaus I, Ross-MacDonald P, Tilford C, Parthasarathy S, Siemers N, Ji R. **Construction of a reference gene association network from multiple profiling data: application to data analysis.** [Bioinformatics](#). 2007 Oct 15;23(20):2716-24.
- Tilford C. **Bio::SeqIO::bsml.** [BioPerl](#), 2001.
- Tilford CA, Kuroda-Kawaguchi T, Skaletsky H, Rozen S, Brown LG, Rosenberg M, McPherson JD, Wylie K, Sekhon M, Kucaba TA, Waterston RH, Page DC. **A physical map of the human Y**

chromosome. [Nature](#) (2001), 409(6822), 943-945.

- Borel CO, Tilford C, Nichols DG, Hanley DF, Traystman RJ. **Diaphragmatic performance during recovery from acute ventilatory failure in Guillain-Barré syndrome and myasthenia gravis.** [Chest](#). 1991 Feb;99(2):444-51.

References

Available on request