



# Benjamin L. Moore

*Bioinformatics PhD and data scientist*

ben@blm.io |  blmoore |  benjaminlmoore

## Currently

I'm a Bioinformatics Scientist at Illumina, based at their R&D site near Cambridge, UK. Here I develop software and use data science, machine learning and statistical programming to analyse large volumes of high-throughput sequencing data.

## Employment

2015– **Illumina, Inc.** Bioinformatics scientist (data scientist and statistical programmer)

## Education

2012–15 **University of Edinburgh** Bioinformatics PhD (supervisors: Colin Semple and Stuart Aitken)  
2011–12 **Imperial College London** MSc Bioinformatics and Theoretical Systems Biology (Distinction)  
2008–11 **University of York** BSc Biology (First class honours)

## Presentations

Tutorials 2014 WikiProject Computational Biology half-day tutorial (ISMB, Boston, USA)

Talks 2015 EdinBR: The Edinburgh R usergroup (various)  
MRC IGMM research seminars and section meetings (various)  
2014 Unravelling higher order chromatin structure (4th Edinburgh Bioinformatics meeting, Edinburgh, UK)  
blogR: pop data analysis and R for the web (Psychology R-users, Edinburgh, UK)

Posters 2015 [Keystone Epigenomics](#) (Keystone CO, USA)  
2014 [ISMB](#) (Boston MA, USA)  
[Genome informatics](#) (Cambridge, UK)  
[Chromatin: From nucleosomes to chromosomes](#) (Cambridge, UK)  
2013 [Genome informatics](#) (CSHL NY, USA)  
[EpiGeneSys](#) (Cambridge, UK)

## Publications

2017 Eberle MA, Fritzilas E, Krusche P, Källberg M, **Moore BL et al.** A reference data set of 5.4 million phased human variants validated by genetic inheritance from sequencing a three-generation 17-member pedigree. [Genome Research](#), 27:157-164.

2016 Yu N, Nützmann H-W, MacDonald JT, **Moore B et al.** Delineation of metabolic gene clusters in plant genomes by chromatin signatures. [Nucleic Acids Research](#), 44(5):2255-2265.

2015 **Moore BL et al.** Integrative modeling reveals the principles of multi-scale chromatin boundary formation in human nuclear organization. [Genome Biology](#), 16:110.  
Fraser J, Ferrai C, Chiariello AM, Schueler M, Rito T, Laudanno G, Barbieri M, **Moore BL et al.** Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. [Molecular systems biology](#), 11:852.

2013 **Moore BL et al.** High-quality protein backbone reconstruction from alpha carbons using Gaussian mixture models. [Journal of computational chemistry](#), 34(22):1881-9.

2011 Gardner PP, Daub J, Tate J, **Moore BL et al.** Rfam: Wikipedia, clans and the “decimal” release. [Nucleic acids research](#), 39(S1):D141-5.

## Non-academic

2014 Pieces in [The Huffington Post](#) (US): Here Are the Most Overrated and Underrated Movies of All Time; Celebrity Twitter Followers, by Gender ([huffingtonpost.com/benjamin-moore](http://huffingtonpost.com/benjamin-moore))

2006– Articles for [English Wikipedia](#): European Nucleotide Archive, RNA thermometer, Toxin-antitoxin system and more ([en.wikipedia.org/wiki/User:Ben\\_Moore](http://en.wikipedia.org/wiki/User:Ben_Moore)).

## Technical skills

R	Python	Data visualisation
Statistics	Machine learning	Mediawiki
LaTeX	Bash	Cluster computing
Git / Mercurial	Emacs	HTML and CSS
UNIX	Inkscape	Javascript

## Projects

blog	I write a data science blog at blm.io where I apply R and Python to open datasets and write-up the results. My posts have been featured by sites including FiveThirtyEight, BuzzFeed, AVclub, io9 and more, and also led to being an invited blogger at the Huffington Post. My original wordpress blog (benjaminlmoore.wordpress.com) has received around 150,000 total views from 180 countries.
EdinbR	I co-founded a usergroup for the R programming language in Edinburgh called EdinbR. I help to organise our monthly meetings and built our website: edinbr.org. Our meetings attract 30-50 statisticians, data scientists and developers for talks and discussion about the R language and its applications. We're sponsored by Revolution Analytics and have an organisational github account at github.com/EdinbR.
dataarea	I entered Imperial College's Summer Data Challenge competition, where entrants analysed a given dataset and then proposed a startup idea based on their results. I developed a modelling technique to select housing areas for investment and was awarded third place: £2,000 and startup support from Imperial Create Lab. My entry can be seen at blm.io/dataarea and the code is shared on my github account.

## Awards

MRC Capacity building scholarship (PhD fees + enhanced stipend)  
MRC Full postgraduate scholarship (MSc fees + stipend)  
ISCB computational biology Wikipedia competition 2013 (doi:10.1371/journal.pcbi.1003242; 2nd place)  
Imperial Create Lab's "Summer data challenge" (summerdatachallenge.com; 3rd place)  
Individual engagement grant from the Wikimedia Foundation to attend ISMB 2014  
Training accreditation for Wikimedia UK, the national chapter of the Wikimedia Foundation  
Front cover design selected for J. Comp. Chem. (34:22) (doi:10.1002/jcc.23388)

## Placements

2010–11	<b>Wellcome Trust Sanger Institute</b> (Cambridge, UK) Worked for two summers on the Rfam database (rfam.xfam.org) in the Bateman group
---------	---

## Links

✉ ben@blm.io	🐦 benjaminlmoore	📖 blmoore
🌐 blmoore	w Ben Moore	🔗 scholar

## References

Available on request.