**EKTA KHURANA**
**CURRICULUM VITAE**
Email: ekk2003@med.cornell.edu
Webpage: http://khuranalab.med.cornell.edu

## RESEARCH EXPERTISE

Computational Biology/Bioinformatics, Genomics, Cancer Genomics, Systems Biology

## PROFESSIONAL EXPERIENCE

**Weill Medical College of Cornell University,** New York, NY 2014 to Present
Assistant Professor (Tenure track)
Meyer Cancer Center
Englander Institute for Precision Medicine
Institute for Computational Biomedicine
Department of Physiology and Biophysics

**Affiliate Member, New York Genome Center** 2018 to Present

**Yale University,** New Haven, CT 2012 to 2014
Associate Research Scientist
Program in Computational Biology and Bioinformatics
Molecular Biophysics and Biochemistry Department

**Yale University,** New Haven, CT 2008 to 2012
Postdoctoral Research Associate
Program in Computational Biology and Bioinformatics
Molecular Biophysics and Biochemistry Department
Advisor: Prof. Mark B. Gerstein

## EDUCATION

**University of Pennsylvania,** Philadelphia, PA 2002 to 2008
Ph. D. in Chemistry (Research area: Computational Biology)
Advisor: Prof. Michael L. Klein
Thesis title: Computational Studies of Natural and Synthetic Ion Channels

**Indian Institute of Technology,** Delhi, India 2000 to 2002
Masters of Science (Research area: Computational Biology)
Advisor: Prof. B. Jayaram
Thesis title: Chemical Model for Genome Analysis

**St. Stephen's College, Delhi University,** Delhi, India 1997 to 2000
Bachelors of Science

## JOURNAL PUBLICATIONS

(* co-first authors, #corresponding author )

1. D Backenroth, Z He, K Kiryluk, V Boeva, L Pethukova, **E Khurana**, A Christiano, J Buxbaum, I Ionita-Laza.
"FUN-LDA: A latent Dirichlet allocation model for predicting tissue-specific functional effects of noncoding variation", American Journal of Human Genetics, 102, 920 (2018)

2. M Bailey ….. **E Khurana** ….. The Cancer Genome Atlas Research Network…..Ding L.
"Comprehensive Characterization of Cancer Driver Genes and Mutations", Cell, 173, 371 (2018)

3. J Kim, FC Geyer, LG Martelotto, CKY Ng, RS Lim, P Selenica, A Li, F Pareja, N Fusco, M Edelweiss, R Kumar, R Gularte-Merida, AN Forbes, **E Khurana**, O Mariani, S Badve, A Vincent-Salomon, L Norton, JS Reis-Filho, B Weigelt
"MYBL1 rearrangements and MYB amplification in breast adenoid cystic carcinomas lacking the MYB-NFIB fusion gene", The Journal of Pathology, 244, 143 (2018)

4. P Dhingra, A Fundichely, A Berger, F Huang, A Forbes, EM Liu, D Liu, A Sboner, P Tamayo, D Rickman[#], M Rubin, **E Khurana**[#]
"Identification of novel prostate cancer drivers using RegNetDriver: A framework for integration of genetic and epigenetic alterations with tissue-specific regulatory network", Genome Biology, 18, 141 (2017)
*Selected for 'Top 10 Papers Reading List' in Regulatory & Systems Genomics by RECOMB/ISCB*

5. A Romanel, S Garritano, B Stringa, M Blattner, D Dalfovo, D Chakravarty, D Soong, K Cotter, G Petris, P Dhingra, P Gasperini, A Cereseto, O Elemento, A Sboner, **E Khurana**, A Inga, M Rubin, F Demichelis
"Inherited determinants of early recurrent somatic mutations in prostate cancer", Nature Communications, 8, 48 (2017)

6. M Feigin, T Garvin, P Bailey, N Waddell, D Chang, D Kelley, S Shuai, S Gallinger, J McPherson, S Grimmond, **E Khurana**, L Stein, A Biankin, M Schatz, D Tuveson
"Recurrent noncoding regulatory mutations in pancreatic ductal adenocarcinoma", Nature Genetics, 49, 825 (2017)

7. T Cuykendall, M Rubin, **E Khurana**[#]
"Non-coding genetic variation in cancer", Current Opinion in Systems Biology, 1, 9 (2017)

8. P Dhingra, Y Fu, M Gerstein[#], **E Khurana**[#]
"Using FunSeq2 for coding and noncoding variant annotation and prioritization", Current Protocols in Bioinformatics, 57, 15.11.1 (2017)

9. **E Khurana**[#],
"Cancer Genomics: Hard-to-reach repairs", Nature, 532, 181 (2016)
*Invited News & Views article*

10. **E Khurana**[#], Y Fu, D Chakravarty, F Demichelis, M Rubin[#], M Gerstein[#]
"Role of non-coding sequence variants in cancer", Nature Reviews Genetics, 17, 93 (2016)

11. The **Cancer Genome Atlas Research Network**
"The molecular taxonomy of primary prostate cancer", Cell, 163(4), 1011 (2015)

12. The **1000 Genomes Project Consortium**
"A global reference for human genetic variation", Nature, 526, 68 (2015)

13. L Lochovsky, J Zhang, Y Fu, **E Khurana**, M Gerstein
"LARVA: An integrative framework for large-scale analysis of recurrent variants in noncoding annotations", Nucleic Acids Research, 43(17), 8123 (2015)

14. Y Fu, Z Liu, S Lu, J Bedford, X Mu, K Yip, **E Khurana**[#], Gerstein M[#]
"FunSeq2: A framework for prioritizing noncoding regulatory variants in cancer", Genome Biology 15, 480 (2014) (co-senior author).

15. K Talbert-Slagle, KE Atkins, KK Yan, **E Khurana**, M Gerstein, EH Bradley, D Berg, AP Galvani, J Townsend,
"Cellular Superspreaders: An Epidemiological Perspective on HIV Infection inside the Body", PLoS Pathogens 10, e1004092 (2014)

16. **E Khurana***, Y Fu*, V Colonna*, X Mu*, HM Kang, …..1000 Genomes Project Consortium….. M Rubin, C Tyler-Smith, M Gerstein,
"Integrative annotation of variants from 1092 humans: application to cancer genomics", Science, 342, 84 (2013)
*Research Highlight in Nature, 502, 144 (2013) and Nature Genetics, 45, 1273 (2013)*

17. **E Khurana***, Y Fu*, J Chen, M Gerstein
"Interpretation of genomic variants using a unified biological network approach", PLoS Computational Biology, 9, e1002886 (2013)

18. **E Khurana**[#],
"Learning to swim in a sea of genomic data", Genome Biology, 14, 315 (2013)
*Invited report on the American Society of Human Genetics meeting, 2013*

19. The **1000 Genomes Project Consortium**
"An integrated map of genetic variation from 1,092 human genomes", Nature, 491, 56 (2012)

20. The **ENCODE Project Consortium**
"An integrated encyclopedia of DNA elements in the human genome", Nature, 489, 57 (2012)

21. L Habegger, S Balasubramanian, D Chen, **E Khurana**, A Sboner, A Harmanci, J Rozowsky, D Clarke, M Snyder, M Gerstein,
"VAT: A computational framework to functionally annotate variants in personal genomes within a cloud-computing environment", Bioinformatics, 28, 2269 (2012)

22. M Gerstein*, A Kundaje*, M Hariharan*, S Landt*, K Yan*, C Cheng*, X Mu*, **E Khurana***, J Rozowsky*, R Alexander*, R Min*, P Alves*, A Abyzov, N Addleman, N Bhardwaj…40 authors…M Snyder,
"Architecture of the human regulatory network derived from ENCODE data", Nature, 489, 91 (2012)

23. D MacArthur…...**E Khurana**…..M Gerstein, C Tyler-Smith,
"A systematic survey of loss-of-function variants in human protein-coding genes", Science, 335, 823 (2012)

24. The **ENCODE Project Consortium**
"A User's Guide to the Encyclopedia of DNA elements", PLoS Biology, 9, e1001046 (2011)

25. R Mills…..**E Khurana**….. J Korbel, 1000 Genomes Project,
"Mapping copy number variation by population-scale genome sequencing", Nature, 470, 59 (2011)

26. Z Lu, K Yip, G Wang, C Shou, L Hillier, **E Khurana**, A Agarwal, R Auerbach, J Rozowsky, C Cheng, M Kato, D Miller, F Slack, M Snyder, R Waterston, V Reinke, M Gerstein,
"Prediction and characterization of non-coding RNAs in C. elegans by integrating conservation,

secondary structure and high throughput sequencing and array data", Genome Research, 21, 276 (2011)

27. **E Khurana[#],** R DeVane, MD Peraro, ML Klein[#],
"Computational study of drug binding to the membrane-bound tetrameric M2 peptide bundle from influenza A virus", Biochimica et Biophysica Acta- Biomembranes, 1808, 530 (2011)

28. M Gerstein….. **E Khurana**….. modENCODE Consortium… R Waterston,
"Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project", Science, 330, 1775 (2010)

29. The **1000 Genomes Project Consortium**,
"A map of human genome variation from population scale sequencing", Nature, 467, 1061 (2010)

30. **E Khurana,** H Lam, C Cheng, N Carriero, P Cayting, M Gerstein,
"Segmental duplications in the human genome reveal details of pseudogene formation", Nucleic Acids Research, 38, 6997 (2010)

31. M Holford, **E Khurana**, K Cheung, M Gerstein,
"Using semantic web rules to reason on an ontology of pseudogenes", Bioinformatics, 26, i71 (2010)

32. Y Arinaminpathy*, **E Khurana*[,#]**, D Engelman, M Gerstein[#],
"Computational analysis of membrane proteins: the largest class of drug targets", Drug Discovery Today, 14, 1130 (2009)

33. Y Liu, D Zheng, S Balasubramanian, N Carriero, **E Khurana**, R Robilotto, M Gerstein,
"Comprehensive analysis of the pseudogenes of glycolytic enzymes in vertebrates: the anomalously high number of GAPDH pseudogenes highlight a recent burst of retrotranspositional activity", BMC Genomics, 10, 480 (2009)

34. H Lam, **E Khurana**, G Fang, P Cayting, N Carriero, K Cheung, M Gerstein,
"Pseudofam: the pseudogene families database", Nucleic Acids Research, 37, D738 (2009)

35. K Talbert-Slagle, S Marlatt, F Barrera, **E Khurana**, J Oates, M Gerstein, D Engelman, A Dixon, D Dimaio,
"Artificial transmembrane oncoproteins smaller than the bovine papillomavirus E5 protein redefine sequence requirements for activation of the platelet derived growth factor β receptor", Journal of Virology, 83, 9773 (2009)

36. **E Khurana[#]**, MD Peraro[#], R DeVane, S Vemparala, WF DeGrado[#], ML Klein,
"Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus", Proceedings of the National Academy of Sciences USA, 106, 1069 (2009)

37. **E Khurana[#]**, R DeVane, A Kohlmeyer, ML Klein,
"Probing peptide nanotube self-assembly at a liquid-liquid interface with coarse-grained molecular dynamics", Nano Letters, 8, 3626 (2008)

38. **E Khurana[#]**, S Nielsen, B Ensing, ML Klein,
"Self-assembling cyclic peptides: molecular dynamics studies of dimers in polar and nonpolar solvents", Journal of Physical Chemistry B, 110, 18965 (2006)

39. **E Khurana**[#], S Nielsen, ML Klein,
"Gemini surfactants at the air/water interface: a fully atomistic molecular dynamics study", Journal of Physical Chemistry B, 110, 22136 (2006)

40. S Dutta, P Singhal, P Agrawal, R Tomer, Kritee, **E Khurana**, B Jayaram,
"A physicochemical model for analyzing DNA sequences", Journal of Chemical Information and Modeling, 46, 78 (2006)

## INVITED CONFERENCE/SYMPOSIA/WORKSHOP LECTURES

1. **Keynote Speaker at HiTSeq18** (High Throughput Sequencing Algorithms track of the 2018 ISMB meeting), Chicago, Illinois, USA (2018)
2. "Non-coding genetic variation in cancer", Gordon Research Conference on Human Genetic Variation and Disease, Maine, USA (2018)
3. "The 3-D cancer genome", Systems Genetics of Cancer, The Francis Crick Institute, London, UK (2017)
4. "Novel prostate cancer drivers identified by integrating genetic and epigenetic alterations with tissue-specific regulatory network", Cancer Genomics Research Network Meetings, New York Genome Center, New York, NY, USA (2017)
5. "Novel prostate cancer drivers identified by integrating genetic and epigenetic alterations with tissue-specific regulatory network", NGS Tech & Applications Congress, Philadelphia, PA, USA (2017)
6. "Novel insights from integrating cancer whole-genomes, epigenomes and transcriptomes with tissue-specific regulatory networks", Systems Genetics of Cancer, Cambridge, UK (2016)
7. "Integrating large-scale genomics data to understand the role of non-coding regions in cancer", Workshop on Models for Oncogenesis, Clonality and Tumor Progression, Mathematical Biosciences Institute, Ohio State University, Columbus, OH (2016)
8. "Tools for analyzing cancer variation", ENCODE User's Meeting, Stanford University, Palo Alto, CA (2016)
9. "Integrating large-scale genomics data to understand the role of non-coding regions in cancer", Early Career Investigators Meeting on Quantitative Problems in Human Health and Genetics, Banff International Research Station, Banff, Canada (2016)
10. Talk: "Integrating large-scale genomics data to understand the role of non-protein-coding regions in cancer" & Panel discussion on "Unleashing the potential of next generation sequencing data for therapeutic development", Festival of Genomics, San Francisco, CA (2015)
11. "Integrating large-scale genomics data to understand the role of non-protein-coding regions in cancer", Next Generation Sequencing Congress, Harvard Medical School, Boston, MA (2015)
12. "Mutations in cancer and precision medicine", Pint of Science festival, New York, NY (2015)
13. "Information in non-coding DNA", American Association for Cancer Research Annual Meeting, San Diego, CA (2014)
14. "Integrative computational models for functional interpretation of genomic sequence variants", Emerging Leaders in Systems-Level Biology Symposium, Cincinnati Children's Hospital Medical Center, Ohio (2014)

## INVITED INSTITUTE/UNIVERSITY TALKS

1. MD Anderson Cancer Center, Houston, Texas, USA (2018)
2. IBM Thomas J. Watson Research Center, Yorktown Heights, NY, USA (2018)
3. Aarhus University, Aarhus, Denmark (2017)
4. Columbia University, New York, NY, USA (2017)
5. Rutgers University, Camden, NJ, USA (2017)
6. Icahn School of Medicine at Mount Sinai, New York, NY, USA (2016)
7. Five Points Lecture, New York Genome Center, New York, NY, USA (2015)
8. University of Massachusetts Medical School, Worcester, MA, USA (2015)
9. Wellcome Trust Sanger Institute, Cambridge, UK (2014)
10. London Research Institute, London, UK (2014)
11. Memorial Sloan Kettering Cancer Center, New York, USA (2014)
12. Institute for Systems Genetics, New York University, New York, USA (2014)

13. Cornell University, Ithaca, USA (2014)
14. University of Toronto, Toronto, Canada (2014)
15. University of Montreal, Montreal, Canada (2014)
16. Cincinnati Children's Hospital Medical Center, Cincinnati, USA (2014)
17. Weill Cornell Medical College, New York, USA (2013)
18. University of Calgary, Calgary, Canada (2013)
19. Jackson Laboratory for Genomic Medicine, Farmington, CT, USA (2013)
20. McGill University and Genome Quebec Innovation Center, Montreal, Canada, (2012)
21. Ste Justine University Hospital Research Center, University of Montreal, Montreal, Canada (2011)
22. National Center for Biological Sciences, Bangalore, India (2009)
23. Yale Center for Genomics and Proteomics, New Haven, CT, USA (2009)
24. Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India (2007)

## ACCEPTED CONFERENCE TALKS

1. "Computational method to identify non-coding cancer drivers", Human Genome Variation Society Meeting, Vancouver, Canada (2016)
2. "Computational methods to identify cancer-driver single nucleotide variants and large rearrangements in non-coding regions", Canadian Computational Biology Conference, University of Toronto, Toronto, Canada (2016)
3. "Computational identification of noncoding cancer drivers from whole-genome sequencing data", RECOMB-CCB (Computational Cancer Biology), Warsaw, Poland (2015)
4. "Computational identification of noncoding cancer drivers from whole-genome sequencing data", Personal Genomes: Discovery, Treatment and Outcomes, Cold Spring Harbor Laboratory, New York (2014)
5. "Identification of non-coding candidate cancer driver mutations using functional annotation of variants from 1,092 humans", Precision Medicine: Personal Genomes and Pharmacogenomics meeting, Cold Spring Harbor Laboratory, New York (2013)
6. "Identification of non-coding candidate cancer driver mutations using functional annotation of variants from 1,092 humans", Cancer Genomics conference, European Molecular Biology Laboratory, Heidelberg, Germany (2013)
7. "Integrative annotation of variants from 1,092 humans: application to cancer genomics", American Society of Human Genetics Meeting, Boston, MA (2013)
8. "Interpretation of genomic variants using a unified biological network approach", The Biology of Genomes meeting, Cold Spring Harbor Laboratory, New York (2013)
9. "Amantadine binding with the Influenza A virus M2 ion channel", CECAM workshop 'Ionic Transport: from Nanopores to Biological Channels', Lyon, France (2007)
10. "Understanding self-assembling cyclic peptide nanotubes in lipid bilayer by molecular dynamics", American Chemical Society National Meeting, San Francisco, CA (2006)
11. "Molecular dynamics study of gemini surfactants at the air/water interface", American Chemical Society National Meeting, San Francisco, CA (2006)
12. "Self-assembly of peptide nanotubes by molecular dynamics study", American Chemical Society National Meeting, San Diego, CA (2005)

## HONORS AND AWARDS

1. Chosen for 'Emerging Leaders in Systems-Level Biology symposium' at Cincinnati Children's Hospital Medical Center (2014)
2. Featured as 'Promising Young Investigator in Genomics' by GenomeWeb (2013)
3. EMBL Corporate Partnership Fellowship for Cancer Genomics conference at EMBL, Germany (2013)
4. Travel Grant for CECAM Workshop 'Ionic Transport: from Nanopores to Biological Channels' at Lyon, France (2007)
5. Chair's fund for Gordon Conference on Computer Aided Drug Design (2007)

6. Marie Curie fellowship awarded by International School of Solid State Physics to attend the course 'Computer Simulations in Condensed Matter' at Erice, Italy (2005)
7. Science Meritorious Award awarded by Delhi University for academic excellence (1997-1998)

## TEACHING AND MENTORING

**Weill Cornell Medical College**

| | |
|---|---|
| 'Tri-I Comp Bio & Medicine graduate program' Steering Committee Member (joint program of WCM, Cornell Ithaca and Memorial Sloan Kettering Cancer Center) | 2016 to Present |
| Co-director of the course 'Tri-I CBM seminar & journal club' | 2016 to Present |
| Thesis/ACE committee member of eleven graduate students from three programs of Weill Cornell Graduate School | 2015 to Present |
| External thesis committee member: | |
| Jiali Zhuang (University of Massachusetts Medical School) | 2015 |
| Malene Juul Rasmussen (Aarhus Univ., Denmark) | 2017 |

**Yale University**

| | |
|---|---|
| Mentored three undergraduate and seven graduate students in Program of Computational Biology and Bioinformatics | 2008 to 2014 |

**University of Pennsylvania,** Dept. of Chemistry

| | |
|---|---|
| Teaching Assistants' Training Workshop leader | 2003 to 2005 |
| Teaching Assistant for undergraduate course General Chem. 101 | 2002 to 2003 |

## OTHER PROFESSIONAL CONTRIBUTIONS/LEADERSHIP

**Leader of 'Genomic Annotations' sub-group** for a collaborative effort between TCGA (The Cancer Genome Atlas) and ICGC (International Cancer Genome Consortium) for pan-cancer analysis of whole genomes (2014-Present)

**Participation in outreach activities of 1000 Genomes and ENCODE consortia**

ENCODE Data Tutorial at American Society of Human Genetics meeting, Boston, MA (2013)
1000 Genomes Data Tutorial at American Society of Human Genetics meeting, San Francisco, CA (2012)

**Invited referee**

Nature, Science, Cell, Nature Genetics, Nature Methods, Nature Communications, Cell Systems, Genome Biology, PLoS Genetics, PLoS Computational Biology, eLife, Bioinformatics, Nucleic Acids Research, Scientific Reports, Modern Pathology, Genome Medicine, Molecular Cancer Research, BMC Systems Biology, BMC Genomics, BMC Cancer, Oncotarget, Biophysical Journal, Journal of Physical Chemistry, Soft Matter, Proteins, PLoS One and BMC Research Notes

**Editorial Board**

Current Protocols in Bioinformatics (2016 to Present)

**Grant Reviewer**

US: NSF BIO Advisory Panel (2016, 2017)
International: European Research Council, Medical Research Council UK, Worldwide Cancer Research (2016)

**Invited workshops**

NHGRI workshop on Computational Genomics & Data Science (2016)

**Program Committee Member**
GLBIO/CCBC-2016 (Canadian Computational Biology Conference)
RECOMB-Seq/CCB-2015, 2017, 2018 (Workshops on Massively Parallel Sequencing and Computational Cancer Biology)

**Conference Abstract Reviewer/Session Moderator**
American Society of Human Genetics Meeting (2016, 2017)

**Last updated: 29 June 2018**