

CV of Yanbin Yin (updated on 1/07/2015)

### Yanbin Yin, Ph.D.

Assistant Professor in Bioinformatics

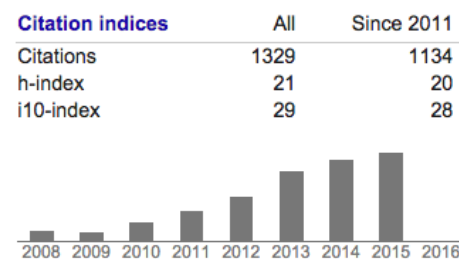
Mailing Address: 325A Montgomery Hall, Department of Biological Sciences, Northern Illinois University, DeKalb, IL 60115

Phone: 815-753-8963;

Email: [yyin@niu.edu](mailto:yyin@niu.edu); [yanbin.yin@gmail.com](mailto:yanbin.yin@gmail.com)

Website: <http://cys.bios.niu.edu/> and <http://www.bios.niu.edu/yin/yin.shtml>

Google Scholar: <https://scholar.google.com/citations?hl=en&user=TgJCXXkAAAAJ>



### Education and training

|                                 |                                       |                 |
|---------------------------------|---------------------------------------|-----------------|
| Yantai University, PR China     | B.S., Biochemistry                    | 9/1996-6/2000   |
| Peking University, PR China     | Ph.D., Biology (Bioinformatics)       | 9/2000-6/2005   |
| University at Buffalo (SUNY)    | Postdoc, Evolutionary Bioinformatics  | 10/2005-10/2007 |
| University of Georgia at Athens | Postdoc, Bioinformatics and Bioenergy | 10/2007-7/2012  |

### Employment history

|                   |  |
|-------------------|--|
| 08/2012 – present | Assistant Professor, Department of Biological Sciences, NIU  |
| 08/2013 – present | Adjunct Assistant Professor, Department of Computer Science, NIU   |
| 01/2011 – 07/2012 | <i>Assistant Research Scientist and group leader</i> , Department of Biochemistry and Molecular Biology and Institute of Bioinformatics, University of Georgia |
| 11/2007 – 01/2011 | <i>Postdoc associate</i> , Department of Biochemistry and Molecular Biology and Institute of Bioinformatics, University of Georgia                             |
| 10/2005 – 10/2007 | <i>Postdoc associate</i> , New York State Center for Excellence in Bioinformatics and Department of Computer Science and Engineering, University at Buffalo    |

### Research interests

- Plant and microbial bioinformatics
- Evolutionary genomics and metagenomics
- Bioenergy related enzymes

### Grants and awards

|           |   |
|-----------|---|
| 2015-2018 | NIH Academic Research Enhancement Award (Grant Number: 1R15GM114706, \$373,400, PI)   |
| 2014      | Research & Artistry Award (\$15,000, PI), Northern Illinois University  |
| 2014      | Travel Award for the 4th Pan-American Congress on Plants and Bioenergy of American Society of Plant Biologists (ASPB), University of Guelph, Canada |
| 2013      | Research & Artistry Award (\$15,000, PI), Northern Illinois University  |

### Research highlights

- I developed an online bioenergy-related enzyme annotation database dbCAN (Nucleic Acids Research 2012, cited 130 times so far), which has been **selected by DOE's Bioenergy Science Centers as one of the Key Advances for Biofuels Production (under Enabling Technologies category in <http://genomicscience.energy.gov/centers/BRCs2014LR.pdf>)**

- I am internationally recognized in the area of bioinformatics analysis for plant cell wall evolution, and was invited to contribute four book chapters including the highly prestigious **Methods in Molecular Biology** series and **Annual Plant Review** series.
- I developed the first computer software (**ORFanFinder**) for finding orphan genes in pathogenic bacteria (funded by NIH) and **HGT-Finder** for finding horizontal gene transfer in pathogenic fungi (Toxins 2015)
- I am the first researcher who demonstrated that the genomic arrangement of bacterial operons is constrained by biological pathways (Proc Natl Acad Sci USA 2010)
- I performed the first genome-wide analysis of orphan genes in viruses (BMC Genomics 2008)

**Most significant peer-reviewed journal publications** # corresponding authors, ^supervised grad students/postdocs, + supervised undergraduate students, \* co-first authors

1. (2016) Hu L, Taujale R<sup>^</sup>, Song J, Guo J, **Yin Y**<sup>#</sup>. Draft genome sequence of *Talaromyces verruculosus* ("Penicillium verruculosum") strain TS63-9, a fungus with great potential for industrial production of polysaccharide-degrading enzymes, **Journal of Biotechnology**, 219:5-6. doi: 10.1016/j.jbiotec.2015.12.017. PMID: 26707807
2. (2015) Nguyen M<sup>+</sup>, Ekstrom A<sup>+</sup>, Li X<sup>^</sup> and **Yin Y**<sup>#</sup>. HGT-Finder: a new tool for horizontal gene transfer finding and application to *Aspergillus* genomes, **Toxins**, 7(10):4035-4053; doi:10.3390/toxins7104035. PubMed PMID: 26473921 [[highly accessed article](#)]
3. (2015) Cary JW<sup>\*</sup>, Han Z<sup>\*</sup>, **Yin Y**<sup>\*</sup>, Lohmar JM, Shantappa S, Harris-Coward PY, Mack B, Ehrlich KC, Wei Q, Arroyo-Manzanares N, Uka V, Vanhaecke L, Bhatnagar D, Yu J, Nierman WC, Johns MA, Sorensen D<sup>+</sup>, Shen H<sup>^</sup>, De Saeger S, Diana Di Mavungu J, Calvo AM. Transcriptome analysis of *Aspergillus flavus* reveals veA-dependent regulation of secondary metabolite gene clusters, including the novel aflavarin cluster. **Eukaryotic Cell** 14:983-997. doi:10.1128/EC.00092-15. PubMed PMID: 26209694 [[Articles of Significant Interest Selected by the Editors](#)]
4. (2015) Taujale R<sup>^</sup>, **Yin Y**<sup>#</sup>. Glycosyltransferase family 43 is also found in early eukaryotes and has three subfamilies in Charophycean green algae, **PLoS ONE**, doi:10.1371/journal.pone.0128409. PubMed PMID: 26023931
5. (2014) Ekstrom A<sup>+</sup>, Taujale R<sup>^</sup>, McGinn N<sup>+</sup>, **Yin Y**<sup>#</sup>. PlantCAZyme: a database for plant carbohydrate-active enzymes, **Database (Oxford)**, DOI: 10.1093/database/bau079. PubMed PMID: 25125445 [[Undergraduate Research and Artistry Day award of NIU 2014](#)]
6. (2014) Cao H<sup>^</sup>, **Yin Y**<sup>#</sup>. Rapid evolution of cellulosome modules by comparative analyses of five Clostridiales genomes. **Bioenergy Research**, 7:1369-1381, DOI 10.1007/s12155-014-9474-0
7. (2014) **Yin Y**<sup>#</sup>, Johns MA, Cao H<sup>^</sup>, Rupani M<sup>^</sup>. A survey of the genomes and transcriptomes of diverse plants and algae reveals new insights into the evolution and function of the cellulose synthase superfamily, **BMC Genomics**, 15:260, DOI: 10.1186/1471-2164-15-260. PubMed PMID: 24708035 [[highly accessed article](#), cited 9 times]
8. (2014) Cao H<sup>^</sup>, Shimura Y, Masanobu K, **Yin Y**<sup>#</sup>. Draft genome sequence of the toxic bloom-forming cyanobacterium *Aphanizomenon flos-aquae* NIES-81. **Genome Announc.** 2(1):e00044-14. doi:10.1128/genomeA.00044-1. PubMed PMID: 24526634 [cited 3 times]
9. (2014) Zhou C, Ma Q, Mao X, **Yin Y**<sup>#</sup>, Xu Y<sup>#</sup>. New Insights Revealed through Comparative Genomic Analysis of 40 Clostridium Genomes, **Bioenergy Research**, 7:1481-1492, doi: 10.1007/s12155-014-9486-9 [cited 3 times]
10. (2013) Zhao Q, Nakashima Q, Chen F, **Yin Y**, Yun J, Shao H, Wang X, and Dixon RA. LACCASE is necessary and non-redundant with PEROXIDASE for lignin polymerization during vascular development in *Arabidopsis thaliana*. **Plant Cell**, 2013 Oct;25(10):3976-87. doi: 10.1105/tpc.113.117770. PubMed PMID: 24143805 [[selected by Faculty 1000](#): <http://f1000.com/prime/718148808>, cited 47 times]

## CV of Yanbin Yin (updated on 1/07/2015)

11. (2013) Ma Q\*, **Yin Y\***, Schell M, Li G, Xu Y. Elucidation of the Dynamic Supercoil Structures of folded E. coli Chromosome: a computational approach, **Nucleic Acids Research**, 2013 Jun 1;41(11):5594-5603. Epub 2013 Apr 17. PubMed PMID: 23599001 [cited 10 times]
12. (2012) Wang S\*, **Yin Y\***, Ma Q, Tang X, Xu Y. Genome-scale identification of cell-wall related genes in Arabidopsis based on co-expression network analysis. **BMC Plant Biology**, 2012, 12:138. PubMed PMID: 22877077 [cited 18 times]
13. (2012) Li G, Köllner TG, **Yin Y**, Jiang Y, Xu Y, Gershenzon J, Pichersky E, Chen F. Nonseed plant *Selaginella moellendorffii* has both seed plant and microbial types of terpene synthases, **Proc Natl Acad Sci USA**, 2012, Sep 4; 109(36):14711-5. PubMed PMID: 22908266 [selected by Faculty 1000: <http://f1000.com/717957718>, listed as must read, cited 25 times]
14. (2012) **Yin Y\***, Mao X\*, Yang JC, Chen X and Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. **Nucleic Acids Res.** web server issue 2012, Jul;40(Web Server issue):W445-51. PubMed PMID: 22645317 [selected by DOE's Bioenergy Science Centers as one of the Key Advances for Biofuels Production, cited 129 times]
15. (2012) Zhang H\*, **Yin Y\***, Olman V, Xu Y. Genomic arrangement of regulons in bacterial genomes. **PLoS One**. 2012;7(1):e29496. Epub 2012 Jan 3. PubMed PMID: 22235300. [cited 10 times]
16. (2011) **Yin Y**, Huang J, Gu X, Bar-Peled M, Xu Y. Evolution of plant nucleotide-sugar interconversion enzymes. **PLoS One**. 2011;6(11):e27995. Epub 2011 Nov 18. PubMed PMID: 22125650. [cited 12 times]
17. (2010) **Yin Y**, Chen H, Hahn MG, Mohnen D, Xu Y. Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. **Plant Physiology** 2010 Aug;153(4):1729-46. Epub 2010 Jun 3. PubMed PMID: 20522722. [cited 47 times]
18. (2010) **Yin Y\***, Zhang H\*, Olman V, Xu Y. Genomic arrangement of bacterial operons is constrained by biological pathways encoded in the genome. **Proc Natl Acad Sci U S A**. 2010 Apr 6;107(14):6310-5. Epub 2010 Mar 22. PubMed PMID: 20308592. [cited 21 times]
19. (2010) Zhao Q, Wang H, **Yin Y**, Xu Y, Chen F, Dixon RA. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. **Proc Natl Acad Sci U S A**. 2010 Aug 10;107(32):14496-501. Epub 2010 Jul 26. PubMed PMID: 20660755; [cited 52 times]
20. (2009) **Yin Y**, Huang J, Xu Y. The cellulose synthase superfamily in fully sequenced plants and algae. **BMC Plant Biol**. 2009 Jul 31;9:99. PubMed PMID: 19646250. [highly accessed article, cited 68 times]
21. (2009) Shen H\*, **Yin Y\***, Chen F, Xu Y, Dixon R. A bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bioenergy production. **Bioenergy Research** 2009;2(4):217-232; [cited 64 times]
22. (2008) **Yin Y**, Fischer D. Identification and investigation of ORFans in the viral world. **BMC Genomics**. 2008 Jan 19;9:24. PubMed PMID: 18205946. [cited 52 times]
23. (2006) **Yin Y**, Fischer D. On the origin of microbial ORFans: quantifying the strength of the evidence for viral lateral transfer. **BMC Evol Biol**. 2006 Aug 16;6:63. PubMed PMID: 16914045 [cited 39 times]

### **Other peer-reviewed journal publications**

24. (2015) Cui J, **Yin Y**, Ma Q, Wang G, Olman V, Zhang Y, Chou WC, Hong CS, Zhang C, Cao S, Mao X, Li Y, Qin S, Zhao S, Jiang J, Hastings P, Li F, Xu Y. Comprehensive Characterization of the Genomic Alterations in Human Gastric Cancer. **Int J Cancer**. 2014 Nov 24. doi: 10.1002/ijc.29352
25. (2014) Zhao Q, Zeng Y, **Yin Y**, Pu Y, Jackson LA, Engle NL, Martin MZ, Tschaplinski TJ, Ding SY, Ragauskas AJ, Dixon RA. Pinoresinol reductase 1 impacts lignin distribution during secondary cell wall biosynthesis in Arabidopsis, **Phytochemistry**. 2014 Aug 5. pii: S0031-9422(14)00278-7. doi: 10.1016/j.phytochem.2014.07.008

26. (2014) Yu J, Jurick WM, Cao H<sup>^</sup>, Yin Y, Gaskins VL, Losada L, Zafar N, Kim M, Bennett JW, Nierman WC. Draft genome sequence of *Penicillium expansum* (R19) that causes postharvest decay of apple fruit. **Genome Announcements**, 2(3):e00635-14. doi:10.1128/genomeA.00635-14
27. (2014) Zhou C, Mao F, **Yin Y**, Huang J, Gogarten JP, et al. AST: An Automated Sequence-Sampling Method for Improving the Taxonomic Diversity of Gene Phylogenetic Trees. **PLoS ONE** 9(6): e98844. doi:10.1371/journal.pone.0098844
28. (2014) Escamilla-Trevino L, Shen H, Hernandez T, **Yin Y**, Xu Y, and Dixon RA. Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (*Panicum virgatum* L.) **Plant Mol Biol**. 2014 Mar;84(4-5):565-76. doi: 10.1007/s11103-013-0152-y. Epub 2013 Nov 5
29. (2013) Liu J, Duan X, Sun J, **Yin Y**, Li G, Wang L, Liu B. Bi-factor Analysis Based on Noise-reduction (BIFANR): A New Algorithm for Detecting Coevolving Amino Acid Sites in Proteins. **PLoS ONE**, 2013 Nov 20;8(11):e79764. doi: 10.1371/journal.pone.0079764
30. (2013) Ma Q, Liu B, Zhou C, **Yin Y**, Li G, Xu Y. An integrated toolkit for accurate prediction and analysis of cis-regulatory motifs at a genome scale. **Bioinformatics**. 2013 Sep 15;29(18):2261-8
31. (2013) Zhang JY, Lee YC, Torres-Jerez I, Wang M, **Yin Y**, Chou WC, He J, Shen H, Srivastava AC, Pennacchio C, Lindquist E, Grimwood J, Schmutz J, Xu Y, Sharma M, Sharma R, Bartley LE, Ronald PC, Saha MC, Dixon RA, Tang Y, Udvardi MK. Development of an integrated transcript sequence database and a gene expression atlas for gene discovery and analysis in switchgrass (*Panicum virgatum* L.). **Plant Journal**. 2013 Apr;74(1):160-73. doi: 10.1111/tbj.12104
32. (2012) Blumer-Schuette SE, Giannone RJ, Zurawski JV, Ozdemir I, Mistry DB, Ma Q, **Yin Y**, Xu Y, Poole FL II, Kataeva I, Adams MWW, Hamilton-Brehm SD, Elkins JG, Larimer FW, Land ML, Hauser L, Cottingham RW, Hettich RL, and Kelly RM. Caldicellulosiruptor core and pan genomes reveal determinants for non-cellulosomal thermophilic deconstruction of lignocellulose. **J of Bacteriology**. 2012 Aug;194(15):4015-28. Epub 2012 May 25.
33. (2012) Mao X, Zhang H, **Yin Y**, Xu Y. The percentage of bacterial genes on leading versus lagging strands is influenced by multiple balancing forces, **Nucleic Acids Res.**, 2012 Sep 1;40(17):8210-8.
34. (2012) Kulkarni AR, Peña MJ, Avci U, Mazumder K, Urbanowicz BR, Pattathil S, **Yin Y**, O'Neill MA, Roberts AW, Hahn MG, Xu Y, Darvill AG, York WS. The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. **Glycobiology**. 2012 Mar;22(3):439-51. Epub 2011 Nov 2. PubMed PMID: 22048859.
35. (2011) Li G, Ma Q, Mao X, **Yin Y**, Zhu X, Xu Y. Integration of sequence-similarity and functional association information can overcome intrinsic problems in orthology mapping across bacterial genomes. **Nucleic Acids Res.** 2011 Dec;39(22):e150. Epub 2011 Sep 29. PubMed PMID: 21965536; PubMed Central PMCID: PMC3239196.
36. (2011) Kong Y, Zhou G, **Yin Y**, Xu Y, Pattathil S, Hahn MG. Molecular analysis of a family of Arabidopsis genes related to galacturonosyltransferases. **Plant Physiol**. 2011 Apr;155(4):1791-805. Epub 2011 Feb 7. PubMed PMID: 21300919; PubMed Central PMCID: PMC3091093.
37. (2011) Dam P, Kataeva I, Yang SJ, Zhou F, **Yin Y**, Chou W, Poole FL 2nd, Westpheling J, Hettich R, Giannone R, Lewis DL, Kelly R, Gilbert HJ, Henrissat B, Xu Y, Adams MW. Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii* DSM 6725. **Nucleic Acids Res.** 2011 Apr;39(8):3240-54. Epub 2011 Jan 11. PubMed PMID: 21227922; PubMed Central PMCID: PMC3082886.
38. (2010) Chou WC, **Yin Y**, Xu Y. GolgiP: prediction of Golgi-resident proteins in plants. **Bioinformatics**. 2010 Oct 1;26(19):2464-5. Epub 2010 Aug 23. PubMed PMID: 20733061; PubMed Central PMCID: PMC2944200.
39. (2010) Zhou C, **Yin Y**, Dam P, Xu Y. Identification of novel proteins involved in plant cell-wall synthesis based on protein-protein interaction data. **J Proteome Res.** 2010 Oct 1;9(10):5025-37. PubMed PMID: 20687615.
40. (2010) Gu X, Glushka J, **Yin Y**, Xu Y, Denny T, Smith J, Jiang Y, Bar-Peled M. Identification of a bifunctional UDP-4-keto-pentose/UDP-xylose synthase in the plant pathogenic bacterium *Ralstonia solanacearum* strain GMI1000, a distinct member of the 4,6-dehydratase and decarboxylase family. **J**

CV of Yanbin Yin (updated on 1/07/2015)

- Biol Chem.** 2010 Mar 19;285(12):9030-40. Epub 2010 Jan 29. PubMed PMID: 20118241; PubMed Central PMCID: PMC2838324.
41. (2010) Yang SJ, Kataeva I, Wiegel J, **Yin Y**, Dam P, Xu Y, Westpheling J, Adams MW. Classification of 'Anaerocellum thermophilum' strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. **Int J Syst Evol Microbiol.** 2010 Sep;60(Pt9):2011-5. Epub 2009 Oct 2. PubMed PMID: 19801388.
  42. (2010) Escamilla-Treviño LL, Shen H, Uppalapati SR, Ray T, Tang Y, Hernandez T, **Yin Y**, Xu Y, Dixon RA. Switchgrass (*Panicum virgatum*) possesses a divergent family of cinnamoyl CoA reductases with distinct biochemical properties. **New Phytol.** 2010 Jan;185(1):143-55. Epub 2009 Sep 15. PubMed PMID: 19761442.
  43. (2009) Mao F, **Yin Y**, Zhou F, Chou WC, Chen H, Zhou C, Dam P, Xu Y. Computational annotation of plant and algae genomes with an emphasis on cell wall biosynthesis. **Bioenergy Research** 2009;2(4): 209-216
  44. (2009) Kong Y, Zhou G, Avci U, Gu X, Jones C, **Yin Y**, Xu Y, Hahn MG. Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in Arabidopsis. **Molecular Plant.** 2009 Sep;2(5):1040-50. Epub 2009 Aug 24. PubMed PMID: 19825678.
  45. (2009) Kataeva IA, Yang SJ, Dam P, Poole FL 2nd, **Yin Y**, Zhou F, Chou WC, Xu Y, Goodwin L, Sims DR, Detter JC, Hauser LJ, Westpheling J, Adams MW. Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "Anaerocellum thermophilum" DSM 6725. **J Bacteriol.** 2009 Jun;191(11):3760-1. Epub 2009 Apr 3. PubMed PMID: 19346307; PubMed Central PMCID: PMC2681903.
  46. (2008) Huang ML, Sun DC, Lou YX, **Yin Y**, Li CY, Zhang Y, Gao G, Wang SQ, Bo XC, Wei LP, Li SG. Analysis and Preparation of Oligonucleotide Microarray of Apoptosis related Genes. **Prog. Biochem. Biophys.** 2008, 35(12): 1451~1460
  47. (2006) Qian Z, **Yin Y**, Zhang Y, Lu L, Li Y, Jiang Y. Genomic characterization of ribitol eichoic acid synthesis in *Staphylococcus aureus*: genes, genomic organization and gene duplication. **BMC Genomics.** 2006 Apr 5;7:74. PubMed PMID: 16595020; PubMed Central PMCID: PMC1458327.
  48. (2006) Huang M, Wang Y, Sun D, Zhu H, **Yin Y**, Zhang W, Yang S, Quan L, Bai J, Wang S, Chen Q, Li S, Xu N. Identification of genes regulated by Wnt/beta-catenin pathway and involved in apoptosis via microarray analysis. **BMC Cancer.** 2006 Sep 7;6:221. PubMed PMID: 16959035; PubMed Central PMCID: PMC1574340
  49. (2005) Chen Y, Zhang Y, **Yin Y**, Gao G, Li S, Jiang Y, Gu X, Luo J. SPD--a web-based secreted protein database. **Nucleic Acids Res.** 2005 Jan 1;33(Database issue):D169-73. PubMed MID: 15608170; PubMed Central PMCID: PMC540047.
  50. (2005) **Yin Y**, Zhang Y, Yu P, Luo JC, Jiang Y, Li SG. Comparative study of apoptosis-related gene loci in human, mouse and rat genomes. **Acta Biochim Biophys Sin.** 2005 May;37(5):341-8. PubMed PMID: 15880263.
  51. (2004) Wang B, Xu Z, Xu C, **Yin Y**, Ding W, Yu H. A Study of Gridifying Scientific Computing Legacy Codes. **Lecture Notes in Computer Science**, 2004;3251, 404 - 412.
  52. (2003) Zhang Y, **Yin Y**, Chen Y, Gao G, Yu P, Luo J, Jiang Y. PCAS--a precomputed proteome annotation database resource. **BMC Genomics.** 2003 Nov 1;4(1):42. PubMed PMID: 14594458; PubMed Central PMCID: PMC293463.
  53. (2003) **Yin Y**, Luo J, Jiang Y. Advances in G-protein-coupled receptors and related bioinformatics study. **Chinese Science bulletin**, 2003;48(6) 511-516.
  54. (2003) Jiang Y, Gao G, Fang G, Gustafson EL, Lavery M, **Yin Y**, Zhang Y, Luo J, Greene JR, Bayne ML, Hedrick JA, Murgolo NJ. PepPat, a pattern-based oligopeptide homology search method and the identification of a novel tachykinin-like peptide. **Mamm Genome.** 2003 May;14(5):341-9. Erratum in: *Mamm Genome.* 2003 Aug;14(8):580. PubMed MID: 12856286.

## **Book publications**

55. (2015) Cao H<sup>^</sup>, Ekstrom A+, **Yin Y#**. Plant carbohydrate active enzyme (CAZyme) repertoires: a comparative study, Book chapter of “Advances in the Understanding of Biological Sciences using Next Generation Sequencing (NGS) Approaches”, **Springer publishing**, 10.1007/978-3-319-17157-9\_8, p115-134
56. (2013) **Yin Y#**. Databases for bioenergy-related enzymes. Book chapter of “Bioenergy Research: Advances & Applications”, **Elsevier Publishing**, edited by Gupta VJ et al., 95-107
57. (2011) **Yin Y**, Mohnen D, Xu Y, Hahn M Glycosyltransferase Family 8 (GT8). Book chapter of “Plant Cell Wall Polysaccharides: Biosynthesis and Bioengineering”, Annual Plant Review series, **Blackwell Publishing**, 2011;41:167-212

## **Submitted publications**

58. Ekstrom A+, **Yin Y#** (2016) ORFanFinder server: towards automated identification and classification of orphan genes, **Nucleic Acids Research**, submitted
59. Wysocki W, Riuz-Sanches E, **Yin Y**, Duvall M (2016) A comparative transcriptomic study between the three main bamboo tribes (Bambusoideae; Poaceae), **Genome Biology and Evolution**, submitted
60. **Yin Y#** (2016) Bioinformatics for analysis of plant cell wall evolution. Invited chapter for “The plant cell wall: Methods and protocols. 2nd edition”, **Methods in Molecular Biology**, submitted

## **Extramural grants**

- **Funded:**  
NIH R15: *Orphan genes in pathogenic bacteria*, Northern Illinois University, 2015-2018 (PI, \$373,400)
- **Unfunded:**  
NSF CAREER: *Evolutionary Genomics of Enzymes for Complex Carbohydrate Metabolism*, Northern Illinois University, 2015-2020 (PI, \$920,743)  
NSF ABI: *Computational tools for lineage-specific genes*, Northern Illinois University, 2014-2017 (PI, \$387,371)  
NSF ABI: *Development of web resources to study the evolution and function of plant cell wall-related genes*, Northern Illinois University, 2013-2016 (PI, \$534,000)  
USDA: *Integrated Research and Education on the Biophysical Basis for Adaptation of Agriculture and Forest Ecosystems to Changing Climate*, Northern Illinois University, 2013-2016 (Co-PI)  
NSF MCB: *Development of web resources to study the evolution and function of plant cell wall-related genes*, Northern Illinois University, 2013-2016 (PI, \$552,637)

## **Teaching experience**

- At NIU: As sole instructor for the following classes: Genetics (BIOS308, 5 credits, ~60 students), Practical Bioinformatics (BIOS441/641, 3 credits, ~20 students), Bioinformatics and Genomics Methods (BIOS443/643, 3 credits, ~20 students), Graduate Seminar (BIOS761, 1 credit, 5-10 students).

| <b>Year</b> | <b>Spring</b> | <b>Fall</b>   |
|-------------|---------------|---|
| 2012        |               | BIOS761E: Next generation sequencing                                      |
| 2013        | BIOS441/641   | BIOS443/643   |
| 2014        | BIOS308       | BIOS441/641, BIOS761E (Plant Seminar): Secondary metabolism gene clusters |

## CV of Yanbin Yin (updated on 1/07/2015)

|      |         |   |
|------|---------|---|
| 2015 | BIOS308 | BIOS441/641, BIOS761A (Microbial Seminar): Horizontal gene transfer in eukaryotes |
| 2016 | BIOS308 |   |

- Before NIU
  - Guest lectures “Computational algorithm for gene finding” (two lectures) in Genomics and Bioinformatics (BCMB3600), University of Georgia, 02/2012
  - Guest lecture “Protein function prediction” in Computational Methods in Bioinformatics (BCMB/BINF8210), University of Georgia, 10/2011
  - Guest lecture “Protein function prediction” in Genomics and Bioinformatics (BCMB3600), University of Georgia, 03/2011
  - Guest lecture “Introduction of Unix commands for biological data processing” in Essential computing skills for biologists (BCMB/BINF4005/6005), University of Georgia, 09/2010
  - Guest lecture “Identification of ORFans in microbial genomes” in Computational Biology (CSE 536), SUNY-buffalo, 03/2006
  - Teaching assistant, Biostatistics, Peking University, 07/2003-01/2004

### **Student mentoring**

*Postdoc*: Huansheng Cao (2012-2014), Elisabeth Fitzek (2014-, part time)

*Graduate students*: Xueqiong Li (BIOS PhD, 2015-), Marcus Nguyen (CS, 2015-), Qasim Rajih (BIOS, 2015-), Rahil Taujale (Bioinfo, 2013-2015), Vamshi Polapally (CS, 2014), Hao Shen (Statistics, 2013-2014), Chenshuang Lu (Statistics, 2014 summer)

*Undergraduate students*: Ashley Purinton (BIOS, 2014-2015), Jessica Morgan (BIOS, 2015), Rhiannon Balazic (BIOS, 2015), Amol Shah (BIOS, 2015), Michael Lewis (BIOS, 2014), Alex Ekstrom (CS, 2013-), Marcus Nguyen (CS, 2014-2015), Blake Fritz (CS, 2014), Emrose Bachhal (BIOS, 2014), Joel Dennison (BIOS, 2013-2014), Tony Pergrossi (CS, 2013), Daniel Sorensen (BIOS, 2013-2014), Nathan McGinn (CS, 2012-2013), Manju Rupani (BIOS, 2012-2013)

### **Professional memberships**

2007- Ad hoc reviewers for NSF MCB, IOS and CAREER programs  
2013- Member, American Association for the Advancement of Science  
2013- Member, American Society for Microbiology  
2013- Member, American Society of Plant Biologists  
2013- Member, Society for Molecular Biology and Evolution  
2010- Associate Editor, BMC Research Notes  
2011- Review Editor, Frontiers in Genetics  
2010-2014 Managing Editor, Frontiers in BioScience

### **Invited/contributed research talks**

1. (3/12/2015) *Plant Cell Wall Evolution: Genomics and Bioinformatics Approaches*. **University of Nebraska**, Center for Biotechnology and School of Biological Sciences
2. (4/27/2015) *Bioinformatics tools to identify orphan genes and horizontally transferred genes in pathogenic microbes*. **U of Illinois at Chicago**, College of Pharmacy at Rockford
3. (5/7/2015) *Plant Cell Wall Evolution: Genomics and Bioinformatics Approaches*. **Southern Illinois University**, Plant Science Department

4. (6/24/2015) *Plant Cell Wall Evolution: Genomics and Bioinformatics Approaches*. **Peking University**, Center for Bioinformatics, College of Life Sciences, China
5. (6/18/2015) *Bioinformatics tools to identify orphan genes and horizontally transferred genes in pathogenic microbes*. **Jilin University**, School of Computer Science, China
6. (6/17/2015) *Bioinformatics tools to identify orphan genes and horizontally transferred genes in pathogenic microbes*. **Jilin University**, School of Basic Medicine, China
7. (6/15/2015) *Plant biomass formation and degradation: bioinformatics and genomics approaches*. **Nankai University**, School of Informatics, China
8. (6/12/2015) *Bioinformatics tools to identify orphan genes and horizontally transferred genes in pathogenic microbes*. **Yangzhou University**, College of Life Sciences, China
9. (6/4/2015) *Plant biomass formation and degradation: bioinformatics and genomics approaches*. **Shandong University**, College of Life Sciences, China
10. (7/17/2013) *Diversity and evolution of cellulosome modules in metagenomes*, the 3rd International Conference on Proteomics & Bioinformatics" (**Proteomics-2013**), DoubleTree by Hilton Philadelphia Center City, USA, July 15-17, 2013
11. (11/15/2012) *Bioinformatics data mining in the era of high-throughput biology*, **SigmaXi NIU Chapter**, Northern Illinois University
12. (9/14/2011) *Bioinformatics study of plant cell wall-related gene families - in support of the bioenergy research within the BioEnergy Science Center of the DOE*, **Brookhaven National Laboratory**
13. (7/19/2011) *dbCAN: a web resource and database for CAZy annotation*, **BESC (BioEnergy Science Center) Science Retreat V**, Chattanooga, Tennessee, July 18-21, 2011
14. (10/27/2010) *Bioinformatics study of plant cell wall biosynthesis*, Integrative BioSystems Institute, **Georgia Tech**, October 27, 2010
15. (7/22/2010) *Evolution and Function of Glycosyltransferase Family 8*, **BESC Science Retreat IV**, Asheville, North Carolina, June 21-23, 2010
16. (12/03/2008) *Computational study of the plant GT8 genes: insight about their origin and function*, **BESC Science Retreat II**, Chattanooga, Tennessee, December 1-3, 2008

#### Conference poster presentations

1. Yin Y, ORFanFinder server: towards automated identification and classification of orphan genes (poster), **Great Lakes Bioinformatics Conference**, Purdue University, IN, May 18-20, 2015
2. Yin Y, PlantCAZyme: a database for plant carbohydrate-active enzymes (poster), **Plants and BioEnergy 2014 Congress**, Conference Center of University of Guelph, Canada June 4 -7, 2014
3. Yin Y, Cao H (2013) Sequence evolution and diversity of cellulosome modules in metagenomes (poster), **Society for Molecular Biology and Evolution (SMBE)** annual meeting
4. Yin Y, Mao X, Yang JC, Chen X and Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. **BioEnergy Science Center retreat VI**
5. Yin Y, Mao X, Yang JC, Chen X and Xu Y (2011) dbCAN: a web resource and database for CAZy annotation. **BioEnergy Science Center retreat V**
6. Yin Y, Chen H, Hahn M, Mohnen D, Xu Y (2010) Evolution and Function of Glycosyltransferase Family 8 (talk), **BioEnergy Science Center retreat IV**
7. Yin Y, Chou WC, Xu Y (2010) The carbohydrate active enzyme space in nature: sub-classification and survey in metagenomes. **BioEnergy Science Center retreat IV**
8. Yin Y, Chou, WC, Xu Y (2010) Application of Phylogenomic Techniques in Studying Glycosyltransferase and Glycoside Hydrolase Families. **DOE Genomic Science 2010 Awardee Workshop VIII and Knowledgebase Workshop**
9. Yin Y, Chou WC, Xu Y (2009) Phylogenetic sub-classification of CAZy families and application in identification of novel glycosyl hydrolases in the environmental metagenomes. **BioEnergy Science Center retreat III**



CV of Yanbin Yin (updated on 1/07/2015)

10. Yin Y, Gu X, Bar-Peled M, Xu Y (2009) Phylogenetic study of NDP-sugar interconversion enzymes: evolutionary relatedness and origin. **BioEnergy Science Center retreat III**
11. Yin Y, Chen H, Cantwell B, Jouligne I, Hahn M, Mohnen D, Xu Y (2009) The plant cell wall biosynthesis related Galacturonosyltransferase (GAUT) and GAUT like (GATL) genes have a different origin than the other Glycosyltransferase family 8 genes. **DOE Genomic Science 2009 Awardee Workshop VII**