Neil D. Clarke

Yale-NUS College
neil.clarke@yale-nus.edu.sg

Professional Positions

Yale-NUS College (Singapore)
2013 –	*Associate Professor, Science*
2013 – 2018	*Head of Studies, Life Sciences (chair)*
2016 – 2019	*Rector, Cendana College*

National University of Singapore (Singapore)
2013 – 2016	*Associate Professor, Department of Biological Sciences*

Genome Institute of Singapore (Singapore)
2005 – 2013	*Senior Group Leader, Computational and Systems Biology*
2005 – 2011	*Deputy Director*

National University of Singapore (Singapore)
2006 – 2013	*Professor (Adjunct), Department of Biochemistry*

Johns Hopkins School of Medicine (Baltimore, MD, USA)
1999 – 2005	*Associate Professor, Department of Biophysics and Biophysical Chemistry*
1992 – 1999	*Assistant Professor, Department of Biophysics and Biophysical Chemistry*

Center for Advanced Research in Biotechnology (Rockville, MD, USA)
1991	*Research Associate*

Johns Hopkins University School of Medicine (Baltimore, MD, USA)
1988 – 1990	*Post-doctoral fellow, Molecular Biology and Genetics and HHMI*

Massachusetts Institute of Technology (Cambridge, MA, USA)
1981 – 1987	*PhD student Department of Biology*

Norwegian Defense Research Institute (Norway)
1980 – 1981	*Visiting Researcher*

Education

1987	**Massachusetts Institute of Technology**. *Ph.D., Biology*
1980	**Yale University** *B.S, Molecular Biophysics and Biochemistry*

Service

Grant and Program Review

National Institutes of Health USA.

2000–04	Biomed. Lib. and Informatics Review Comm. (bioinformatics study section)
2003	ENCODE review panel, National Human Genome Research Institute
2002	Special Emphasis Panel: Bioterrorism-Related Research
2000	Special Emphasis Panel: Protein Information Resource

Department of Energy USA.

1999	Biotechnology Research Advisory Committee, Lawrence-Livermore Labs
1996	Reviewer of National Laboratory programs in DNA repair

American Cancer Society.

1994	Advisory Committee on Nucleic Acids and Protein Synthesis

Community-wide experiments in blinded prediction

CASP (protein structure; http://predictioncenter.org)

1994	Invited participant in inaugural CASP competition/experiment
2006	Lead assessor for 'template-free' (new fold) structure prediction, CASP7

DREAM (systems biology; http://collections.plos.org/dream

2007	Participant in inaugural DREAM challenge
2008	Posed and judged a gene expression prediction challenge for DREAM3

Journals

2011–13	**Editorial Board**, Algal Research.
continuous	**Reviewer**: Biochemistry, Bioinformatics, BMC Bioinformatics, BMC Genomics, EMBO Journal, Genome Research, Journal of the American Chemical Society, Journal of Biological Chemistry, Journal of Bioinformatics and .Computational Biology, Molecular Biology and Evolution, Molecular Systems Biology, Nature Biotechnology, Nature Structural Biology, Nucleic Acids Research, PLoS Computational Biology, PLoS ONE, Proceedings of the National Academy of Science, Protein Science, Proteins, Science, Scientific Reports

Lectures and Conference Talks

2013 **NGS (Next Generation Sequencing) Asia Congress** *Singapore*

2013 **NYU-Abu Dhabi – Center for Genomics and Systems Biology** *Abu Dhabi, UAE*

2013 **National University of Singapore, Biological Sciences** *Singapore*

2012 **Protein Society** *San Diego, CA*

2012 **4th International Singapore Lipid Symposium** *Singapore*

2012 **Alternative Aviation Fuel in Asia & ASEAN Algae Biofuel Initiative Conference** *Singapore*

2011 **Economic Development Board 50th Anniversary Learning Fest: Urbanization and Sustainability** *Singapore*

2011 **Singapore Centre for Environmental and Life Science Engineering, Nanyang Technological University** *Singapore*

2011 **1st International Conference on Algae Biomass, Bioproducts and Biofuels** *St. Louis, MO*

2011 **Complexity and Systems Biology of Microbial Biofuels** *Warwick U., England*

2011 **Biofuels (Keystone Conference)** *Singapore*

2010 **Algae World Asia** *Singapore.*

2010 **4th Annual Algae Biomass Summit** *Phoenix, AZ*

2010 **ASTAR-NKTH Symposium** *Budapest, Hungary*

2010 **Singapore Lipid Symposium** *Singapore*

2009 **Chromatin: Histones, Nucleosomes, Chromosomes and Genomes** *Singapore*

2008 **RECOMB Regulatory Genomics** *Broad Institute, Cambridge, MA*

2008 **Dialogue for Reverse Engineering Assessments and Methods** *Broad Institute, Cambridge, MA*

2008 **CASP 7.5 (Critical Assessment of Structure Prediction)** *Madrid*

2007 **Dialogue for Reverse Engineering Assessments and Methods** *New York*

2007 **University of California, Los Angeles**

2007 **Systems Biology: Genomic Approaches to Transcriptional Regulation** *Cold Spring Harbor, NY*

2007 **Nanyang Technological University** *Singapore*

2006 **Critical Assessment of Structure Prediction** *Asilomar, CA*

2006 **Temasek Life Sciences Laboratory** *Singapore*

2005 **Jackson Labs/TIGR 8th Annual Conference on Computational Genomics** *Boston,MA*

2005 **Yale University**

2005 **Columbia University**

Publications

Reviews

Clarke ND. (2010) Protein engineering for bioenergy and biomass-based chemicals. ***Curr. Opin. Struct. Biol.*** (2010) 20:527-532

Lieb JD and **Clarke ND** (2005) Control of transcription through intragenic patterns of nucleosome composition ***Cell*** 123: 1187-1190

Desjarlais J and **Clarke ND**. (1998) Computer search algorithms in protein design and modification. ***Curr. Opin. Struct. Biol.*** 8:471-475

Clarke ND. (1995)Sequence 'minimization': exploring the sequence landscape with simplified sequences. ***Curr. Opin. Biotech.*** 6: 467-472.

Research

One Thousand Plant Transcriptomes Initiative*. (2019) A Phylogenomic View of Evolutionary Complexity across Green Plants. *Nature*. 574: 679–685 *[193 authors, most alphabetical]

Goh FQY, Jeyakani J, Tipthara P, Cazenave-Gassiot A, Ghosh R, Bogard N, Yeo X, Wong G-KS, Melkonian M, Wenk MR, and **Clarke ND**. (2019). Gains and losses of metabolic function inferred from a phylotranscriptomic analysis of algae. ***Scientific Reports***. *9*(1), 10482.

Goh FQY, Jeyakani J, Cazenave-Gassiot A, Tipthara P, Wenk M, and **Clarke, ND**. (2017) Core features of triacylglyceride production in *Ettlia oleoabundans* revealed by lipidomic and gene expression profiling under distinct induction conditions. ***Algal Research***. 26: 453-462

Aow JSZ, Xue X, Run J-Q, Lim GFS, Goh WS, and **Clarke ND**. (2013). Differential binding of the related transcription factors Pho4 and Cbf1 can tune the sensitivity of promoters to different levels of an induction signal. ***Nucleic Acids Research***.

Wollmann H, Holec S, Alden K, **Clarke ND**, Jacques P-É, and Berger F. (2012) Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the Arabidopsis transcriptome. ***PLoS Genetics*** 8: e1002658

Johnson MTJ, …[6 authors], **Clarke ND**, [29 authors]. (2012) Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. ***PLoS ONE***, 7: e50226

Song L, …[19 authors], **Clarke ND**, Birney E, Iyer, VR, Crawford, GE, Lieb JD, and Furey, TS. (2011) Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. ***Genome Research*** (doi:10.1101/gr.121541.111)

Joseph R, Orlov YL, Huss M, Sun W, Kong SL, Ukil L, Pan YF, Li G, Lim M, Thomsen JS, Ruan Y, **Clarke ND**, Prabhakar S, Cheung E, Liu ET. (2010) Integrative model of genomic factors for determining binding site selection by estrogen receptor-α. *Molecular Systems Biology* vol. 6 pp. 456

Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, Huss M, Soh BS, Kraus P, Li P, Lufkin T, Lim B, **Clarke ND**, Bard F and Ng HH. (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* epub Oct 17 2010

Guo G, Huss M, Tong GQ, Wang C, Sun LL, **Clarke ND**, Robson P (2010) Resolution of cell fate decisions revealed by single cell gene expression analysis from zygote to blastocyst *Developmental Cell* 18(4):675-85

Goh WS, Orlov Y, Li J, **Clarke ND** (2010) Blurring of high-resolution data shows that the effect of intrinsic nucleosome occupancy on transcription factor binding is mostly regional, not local *PLOS Comp Biol* 6 : e1000649.

Clarke ND, Bourque G (2010) Success in the DREAM3 Signaling Response Challenge Using Simple Weighted-Average Imputation: Lessons for Community-Wide Experiments in Systems Biology *PLOS ONE* 5 e8417

Prill RJ, Marbach M, Saez-Rodriguez J, Sorger PK, Alexopoulos LG, Xue X, **Clarke ND**, Altan-Bonnet G and Stolovitzky G (2010) Towards a Rigorous Assessment of Systems Biology Models: The DREAM3 Challenges *PLOS ONE* 5 e8417

Liu X and **Clarke ND** (2009). Transcription Factor Binding Probabilities in Orthologous Promoters: An Alignment-Free Approach to the Inference of Functional Regulatory Targets. In: Ciccarelli F.D., Miklós I. (eds) Comparative Genomics. RECOMB-CG 2009, Springer. **Lecture Notes in Computer Science** 5817**:**229-240

Vega V, Woo X, Hamidi H, Yeo H, Yeo Z, Bourque G and **Clarke ND**. (2009) Inferring direct regulatory targets of a transcription factor in the DREAM2 challenge. *Annals NY Acad Sci* 1158 215-223

Yeo ZX, Yeo HC, Yeo, JKS, Yeo AL, Li Y, **Clarke ND**. (2009) Inferring transcription factor targets from gene expression changes and predicted promoter occupancy *J Comp Bio* 16 357-368

Badis G, Chan ET, van Bakel H, Pena-Castillo L, Tillo D, Tsui K, Carlson CD, Gossett AJ, Hasinoff MJ, Warren CL, Gebbia M, Taludker S, Yang A, Mnaimneh S, Terterov D, Coburn D, Yeo AL, Yeo ZX, **Clarke ND**, Lieb JD, Ansari AZ, Nislow C, Hughes TR. (2008) A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Molecular Cell* 32 878-887

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, **Clarke ND**, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells *Cell* 133 1106-1117.

Jauch R, Yeo HC, Kolatkar PR, **Clarke ND** (2007) Assessment of CASP7 structure predictions for template free targets. ***Proteins*** 2007, 69 Suppl 8:57-67

Izarzugaza JM, Grana O, Tress ML, Valencia A, **Clarke ND** (2007) Assessment of intramolecular contact predictions for CASP7. ***Proteins*** 69 Suppl 8:152-158.

Clarke ND, Ezkurdia I, Kopp J, Read RJ, Schwede T, Tress M (2007) Domain definition and target classification for CASP7. ***Proteins*** 69 Suppl 8:10-18.

Battey JN, Kopp J, Bordoli L, Read RJ, **Clarke ND**, Schwede T (2007) Automated server predictions in CASP7. Proteins 2007, 69 Suppl 8:68-82.

Liu X, Lee C-K, Granek JA, **Clarke ND** and Lieb, JD (2006) Whole genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. ***Genome Research*** 16 1517-1528

Tang L, Liu X and **Clarke ND** (2006) Inferring direct regulatory targets from expression and genome location analyses: a comparison of transcription factor deletion and overexpression ***BMC Genomics*** 7 215

Granek JA and **Clarke ND**. (2005) Explicit equilibrium modeling of transcription factor binding and gene regulation. ***Genome Biology*** 6 R87.

Liu X, Noll DM, Lieb JD and **Clarke ND**. (2005) DIP-ChIP: Rapid and accurate determination of DNA binding specificity. ***Genome Research*** 15 421-427

Carroll K, Pradhan D, Granek JA, **Clarke ND** and Corden J. (2004) Identification of cis-elements directing termination of yeast non-polyadenylated snoRNA transcripts. ***Mol. Cell Bio***. 24 6241-6252

Daniels DS, Woo TT, Luu KX, Noll DM, **Clarke ND**, Pegg AE and Tainer JA (2004) DNA binding and nucleotide flipping by the human DNA repair protein AGT ***Nature Struct. Mol. Biol***. 11 714-720

Clarke ND and Granek JA. (2003) Rank order metrics for quantifying the association of sequence features with gene regulation. ***Bioinformatics*** 19 212-218

Liu X. and **Clarke ND**. (2002) Rationalization of gene regulation by a eukaryotic transcription factor: Calculation of regulatory region occupancy from predicted binding affinities ***J. Mol Biol***. 323 1-8

Noll DM and **Clarke ND** (2001) Covalent capture of a human O(6)-alkylguanine alkyltransferase-DNA complex using N1,O6-ethanoxanthosine, a mechanism-based crosslinker. ***Nucl. Acids Res***.. 29 4025-4034

Brown LR, Deng J, and **Clarke ND**. (2000) Dominant sensitization variants of human O6-methylguanine-DNA-methyltransferase obtained by a mutational screen of surface residues. ***Mutation Research*** 459 81-87

Gogos A, Jantz D, Sentürker S, Richardson D, Dizdaroglu M, and **Clarke ND**. (2000) Assignment of enzyme substrate specificity by principal component analysis of aligned protein sequences: an experimental test using DNA glycosylase homologs. ***Proteins*** 40: 98-105

Gogos A and **Clarke ND**. (1999) Characterization of an 8-oxoguanine DNA glycosylase from Methanococcus jannaschii. ***J. Biol. Chem***.274 30447-30550

Noll DM, Gogos A, Granek JA, and **Clarke ND**. (1999) The C-terminal domain of the adenine-DNA glycosylase MutY confers specificity for 8-oxoguanine•adenine mispairs and may have evolved from MutT, an 8-oxo-dGTPase. ***Biochemistry*** 38 6374-6379

Clarke ND and Berg JM. (1998) Zinc fingers in C. elegans: Finding families and probing pathways ***Science*** 282 2018-2022

Yuan S-M. and **Clarke ND**. (1998) A hybrid sequence approach to the Parcelsus Challenge. ***Proteins*** 30 136-143.

Paalman SF, Sung C, and **Clarke ND** (1997) Specificity of DNA repair methyltransferases determined by competitive inactivation with oligonucleotide substrates: evidence that *Escherichia coli* Ada repairs O6-methylguanine and O4-methylthymine with similar efficiency. ***Biochemistry*** 36 11118

Paalman SR, Noll D, and **Clarke ND** (1997) Formation of a covalent complex between methylguanine methyltransferase and DNA via disulfide bond formation between the active site cysteine and a thiol containing analog of guanine. ***Nucleic Acids Res***. 25 1795

Gogos A, Cillo J, **Clarke ND**, and Lu A-L. (1996) Specific recognition of A/G and A/7,8-dihydro-8-oxoguanine (8-oxoG) mismatches by Escherichia coli MutY: Removal of the C-terminal domain preferentially affects A/8-oxoG recognition. ***Biochemistry*** 35 16665-16671.

Clarke ND. (1995) Co-variation of residues in the homeodomain sequence family. ***Protein Science*** 4 2269-2278

Klemba M, Gardner KH, Marino S, **Clarke ND**, and Regan L. (1995) Novel metal-binding proteins by design. ***Nature: Structural Biology*** 2 368-373

Clarke ND and Yuan S-M. (1995) Metal Search: A computer program that helps design tetrahedral metal binding sites. ***Proteins*** 23 256-263

Clarke ND, Kissinger CR, Desjarlais J, Gilliland GL, and Pabo CO. (1994) Structural studies of the engrailed homeodomain. ***Protein Science*** 3 1779-1787

Clarke ND. (1994) A proposed mechanism for the self-splicing of proteins. ***Proc Natl Acad Sci USA*** 91 11084-11088

Shortle D and **Clarke ND**. (1993) Alpha-helix propensity of amino acids [technical comment] **Science** 262 917.

Clarke ND, Beamer LJ, Goldberg HR, Berkower C, and Pabo CO. (1991). The DNA binding arm of λ repressor: critical contacts from a flexible region *Science* 254 267-270

Regan LJ and **Clarke ND**. (1990). A tetrahedral Zn(II)-binding site introduced into a designed protein. *Biochemistry* <u>29</u>, 10878-10883

Bowie JU, **Clarke ND**, Pabo CO, and Sauer RT. (1990) Identification of protein folds: matching hydrophobicity patterns of sequence sets with solvent accessibility patterns of known structures. *Proteins* <u>7</u>, 257-264

Clarke ND, Lien DC, and Schimmel P. (1988). Evidence from cassette mutagenesis for a structure-function motif in a protein of unknown structure. *Science* <u>240</u>, 521-523.

Clarke ND, Kvaal M, and Seeberg E (1984) Cloning of Escherichia coli genes encoding 3-methyladenine DNA glycosylases I and II. Mol. Gen. Genet. 197: 368–372.

Sancar A, **Clarke ND**, Griswold J, Kennedy WJ, and Rupp WD (1981) Identification of the uvrB gene product. *J Mol Biol* 148: 63–76

Sancar A, Wharton RP, Seltzer S, Kacinski BM, **Clarke ND**, and Rupp WD. (1981) Identification of the uvrA gene product. J Mol Biol 148: 45–62