

Detection of DNase-I Hypersensitive sites at the human KIBRA gene locus by Quantitative PCR

Krupa Arun Navalkar

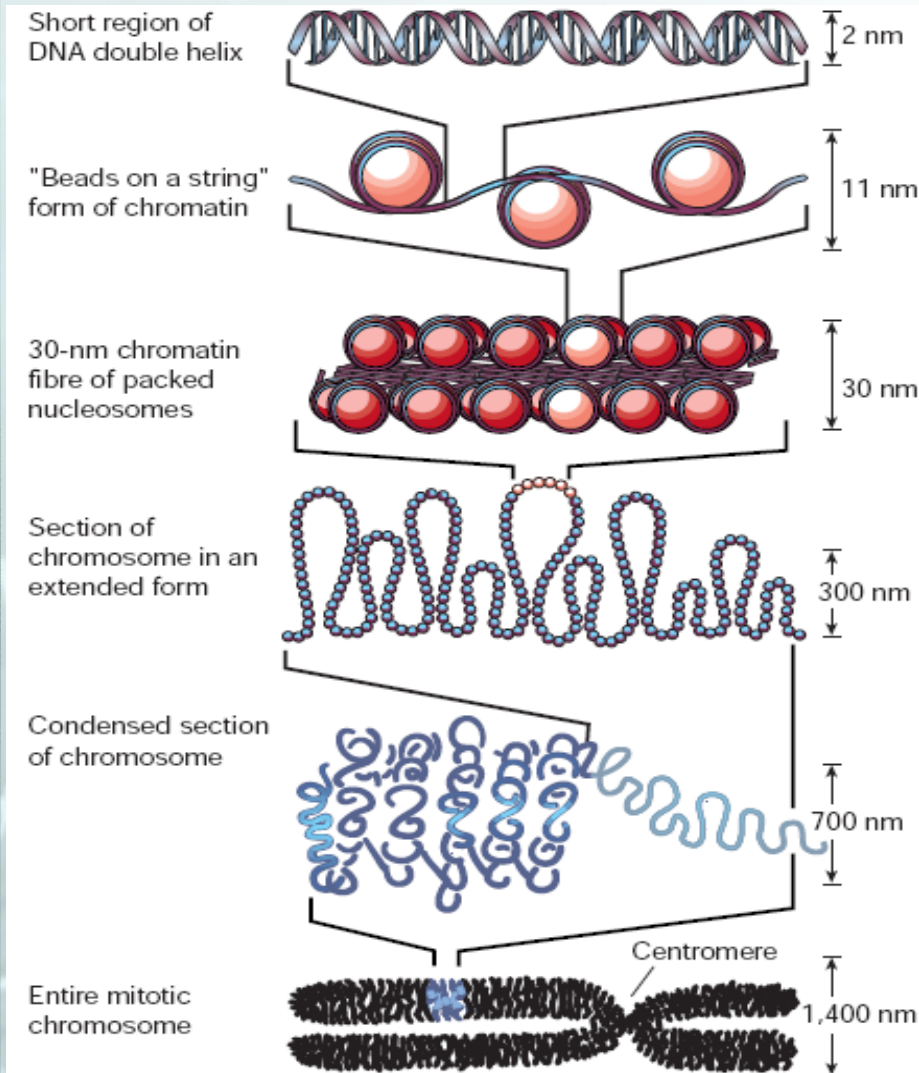
Internship Advisor: Dr. Jeffrey Touchman

PSM Computational Biosciences

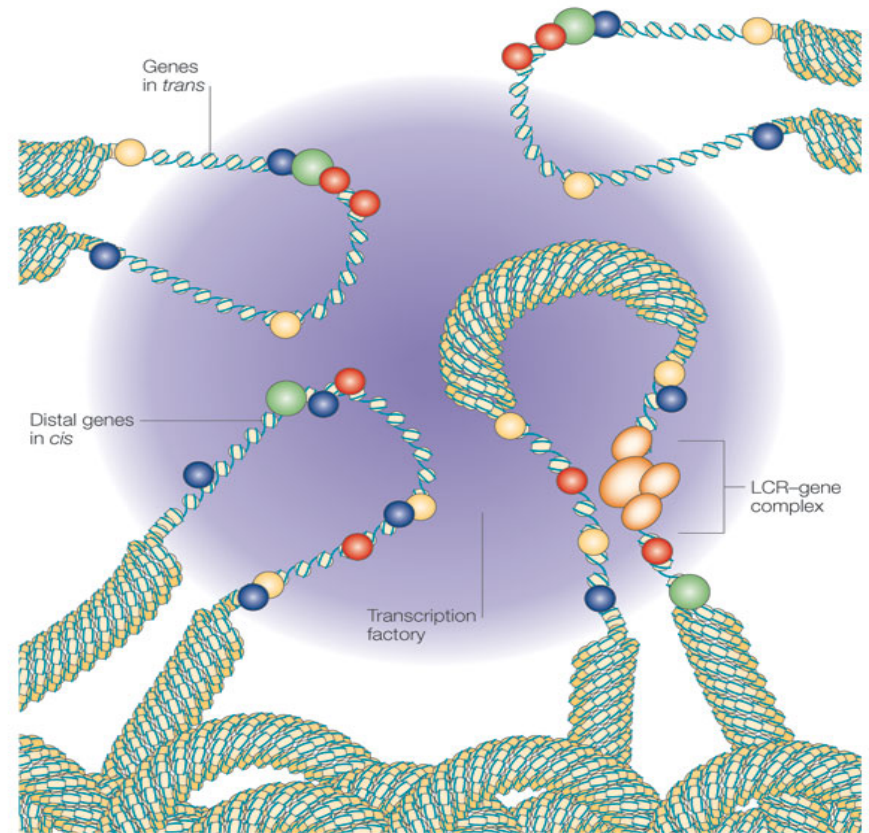
Arizona State University

28th July, 2008

What are we looking for?



Ref: Alberts, B. et al. (1998) Essential Cell Biology: An Introduction to the Molecular Biology of the Cell. Garland, New York.



Copyright © 2005 Nature Publishing Group
Nature Reviews | Genetics

Ref: FIGURE 5: Transcription factory model

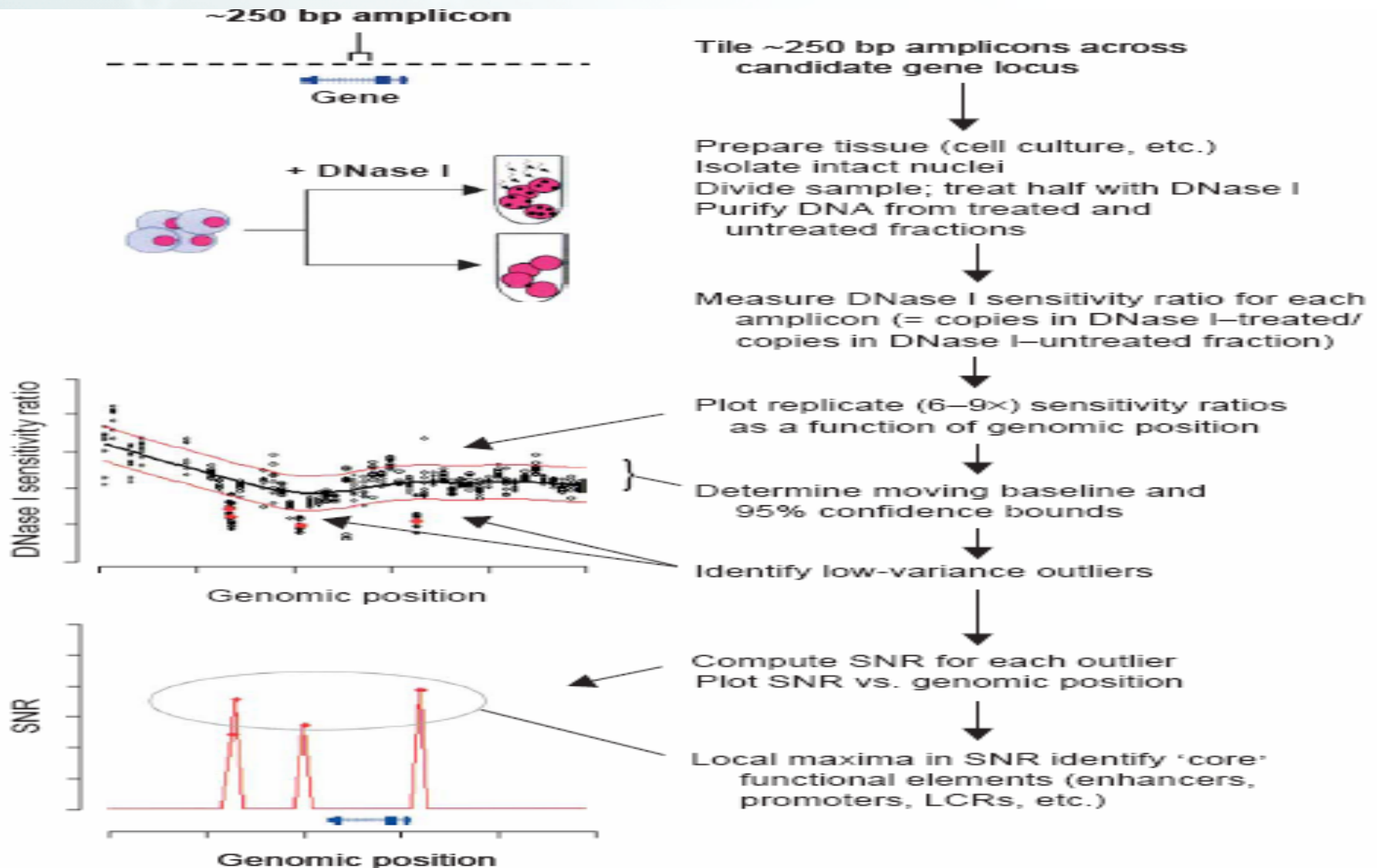
From the following article:

Lyubomira Chakalova, Emmanuel Debrand, Jennifer A. Mitchell, Cameron S. Osborne & Peter Fraser (2005) Replication and transcription: Shaping the landscape of the genome. Nature Reviews Genetics 6, 669-677.

Why the Kibra gene?

- Human Chromosome no. 5: 5q34-q35.2 Total Region: 167,769,199 - 167,801,644 (32.4 kb)
- 1,113 amino acid long protein, molecular mass; 125 kD
- Precisely localize cis-regulatory sequences and other functional modalities encoded by DNase-I Hypersensitive sites (traditionally localized using Southern blot)
- To identify functional, non-coding elements that regulate transcription so as to understand disease processes associated with the Kibra.
- It is expressed in breast cancer cells in relation to its interaction with the Dynein Light Chain 1(DLC1) protein leading to growth stimulation in the same.
- It is also found to be expressed in its truncated form, where its missing the first 223 amino acids in memory retrieval related brain structures such as the hippocampus and temporal lobes, and functional magnetic resonance imaging (fMRI) studies show a significant role of Kibra in memory performance.

Detection of DNase-I hypersensitive sites



Ref: Dorschner et. al. 2004

Components of the system

- Primers: contiguous or minimally overlapping ~ 250 bp, target amplicon size 250 bp, primer Tm ($60^{\circ}\text{C} \pm 2^{\circ}\text{C}$), length 18-23, designed using program Primer3

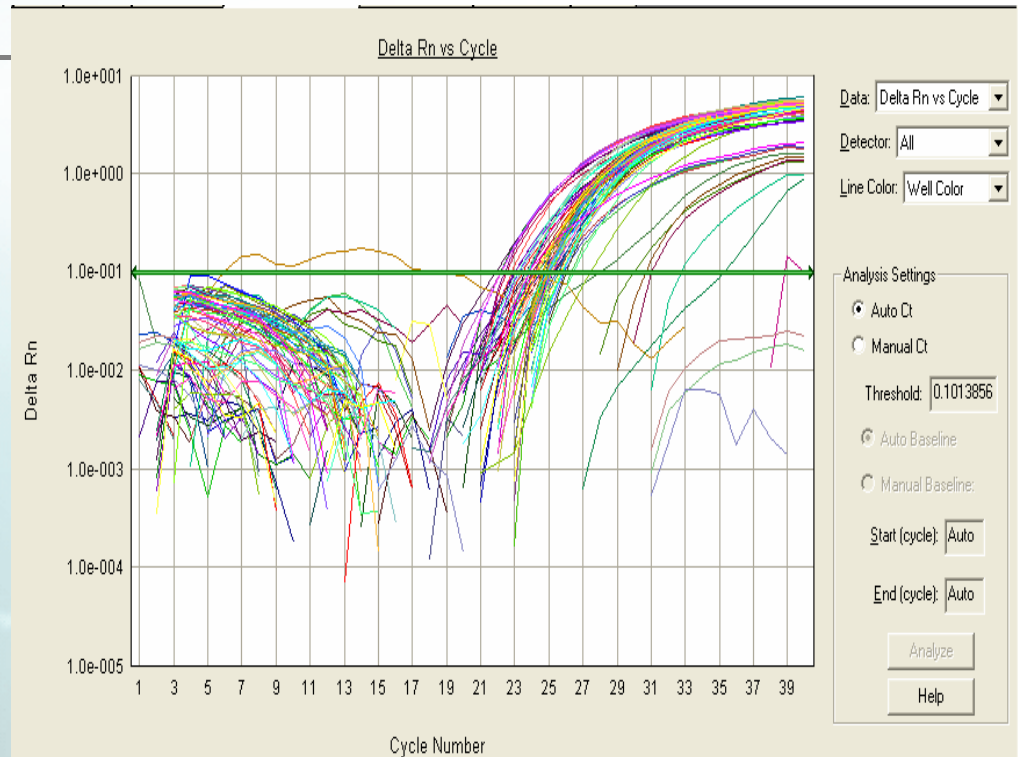
- Quantitative PCR: Ct values and Efficiencies

(10-80 U/ml) DNase –I treated and untreated 30 primers spanning the region of interest (chr5:167,769,118-167,801,709) within the Kibra gene (eight replicates each).

(10-80 U/ml) DNase – I treated and untreated RHO primer (eight replicates each)

- Standard Curve for RHO primer with input amounts of DNA: 300 ng, 30 ng, 3 ng, 0.3 ng, 30 pg, 3 pg

Amplification Plot as seen in ABI Prism 7900 SDS software



IMR-90 is a normal lung fibroblast, U-87 MG (Glioblastoma, brain): low expresser of Kibra

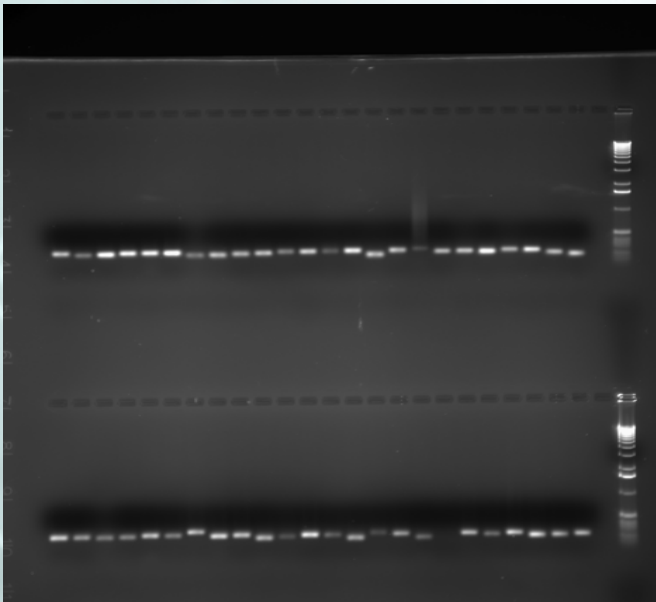
T-47D (breast cancer), MDA-MB-436 (breast tissue): high expressers of Kibra.

Primary Data Analysis

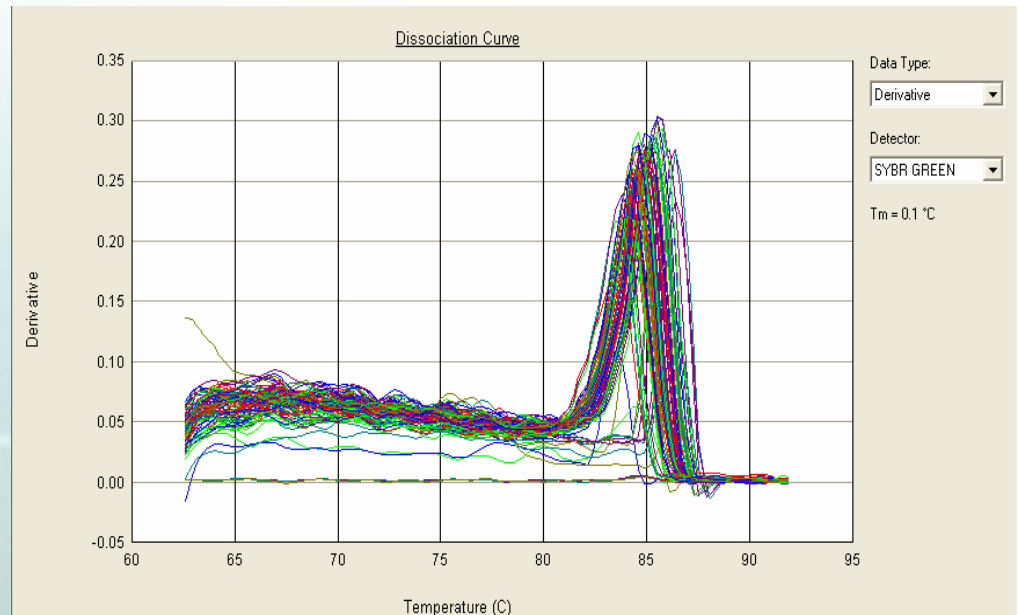
- Normalized fluorescence data are obtained
- An amplification curve and nth-order polynomial fit were computed for each reaction.
- The Ct values were then determined for each curve.
- Melting curve analysis will be done for each amplicon to discard those yielding multiple products.
- The amplification efficiency of a reference amplicon selected from the inactive and DNase-insensitive RHO (Rhodopsin) locus (3q21-q24) was determined empirically for every reaction plate using a standard dilution series of DNA. This is an internal quality control for every plate.
- Derive the efficiency of each test amplicon from the amplification curve using LinRegPCR software.

Quality Control for the Primers

Quality Control for Primer rows A-D



Dissociation Curve as seen in ABI Prism 7900 SDS software

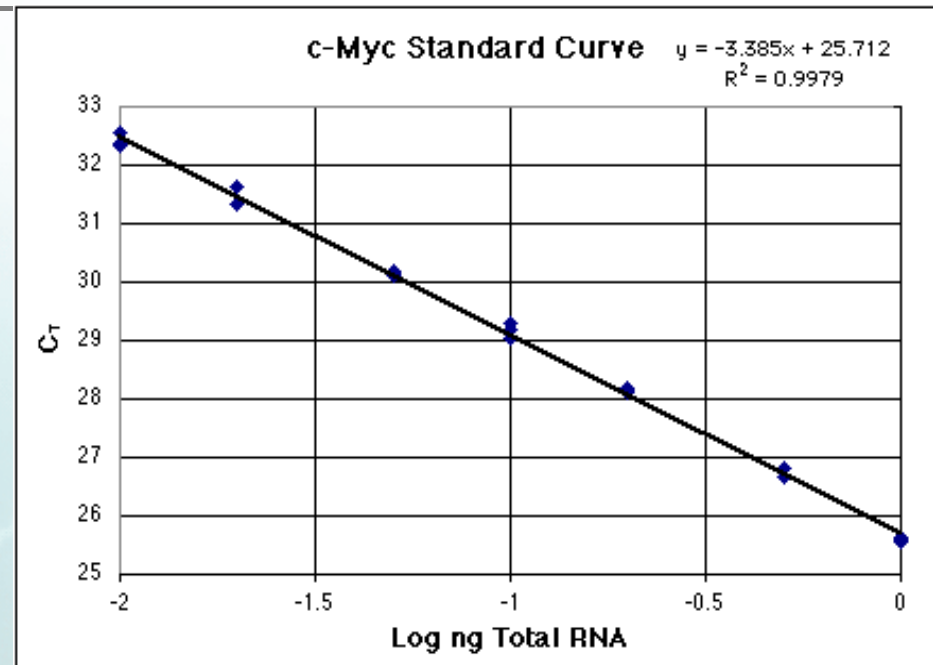


In the above picture all the PCR amplicons were selected in the software interface which is why multiple melting point curves are seen. Upon selecting individual PCR amplicons we can clearly see that each amplicon produces only a single product thereby adding information to the already existing gel electrophoresis image quality control data, stating the same.

A Typical Q-PCR Reaction...

- Perform a Standard Curve Experiment to validate that the efficiency of the test primer and the internal reference gene primer are the same.
- This is a graph of Ct values vs. Log input amount at different concentrations of input RNA/DNA.
- PCR Efficiency is the rate at which a PCR amplicon is generated, commonly expressed as a percentage value.
- If a particular PCR amplicon doubles in quantity during the geometric phase of its PCR amplification then the PCR assay is said to have 100% efficiency.

Ref: User Bulletin 2: ABI PRISM 7700 SDS (1997)



Determine the efficiency of this amplicon using the equation

$$E = 10^{-1/\text{slope}} \text{ (i.e. } 10^{-1/-3.32} = 2)$$

A Standard Curve Slope of -3.32 indicates reactions with 100% efficiency

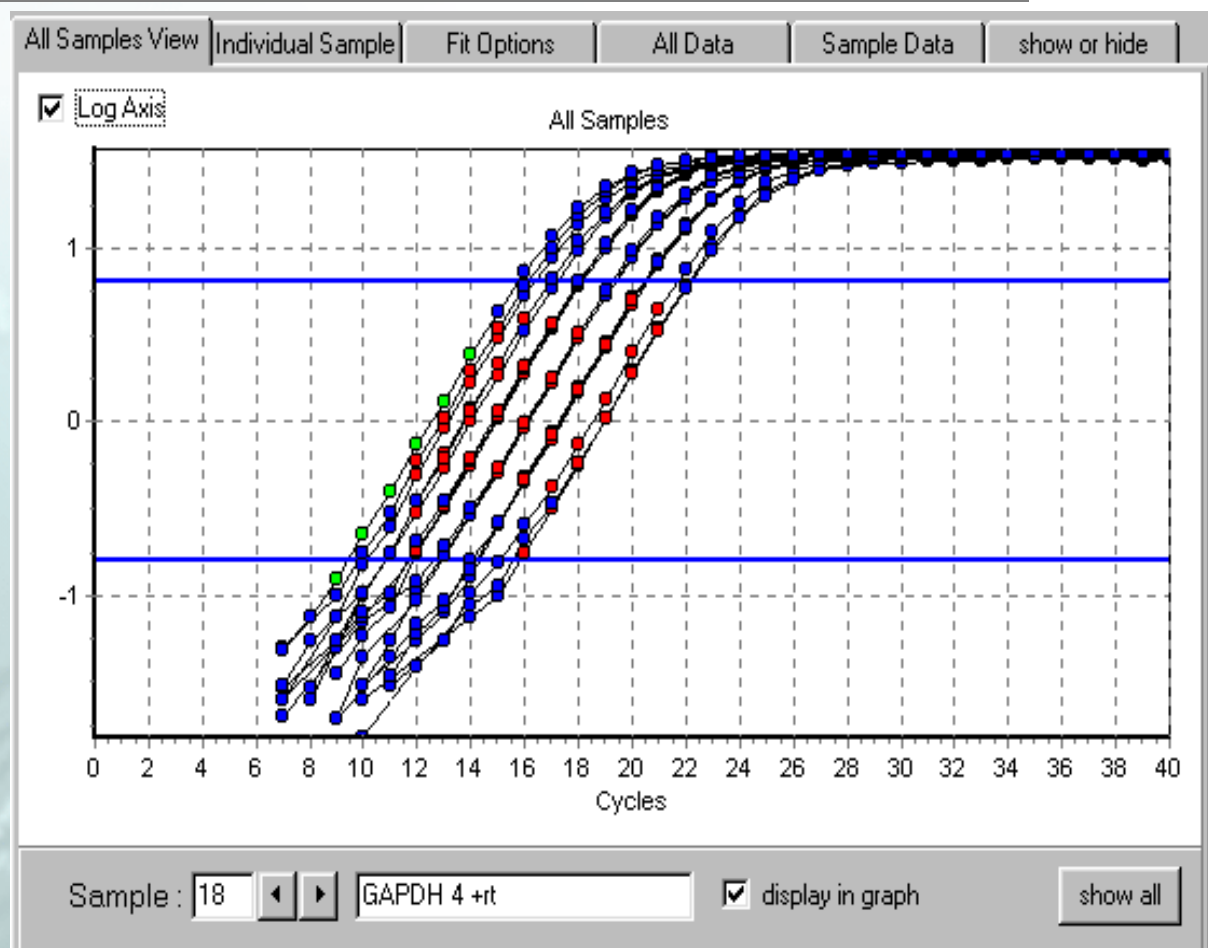
LinRegPCR – Ruijter et al – Assumption free analysis of Q-PCR

Assumption: $E=2$ for target and reference.

In practice, most efficiencies vary from 1.8 – 2.0 which shows up as a 10-fold difference.

Amplification Efficiency Correction criteria:

- The efficiency values must be within the median range of efficiency values ± 0.05
- The range of R values is between 0.995 to 0.999 and,
- The range of



LinRegPCR...

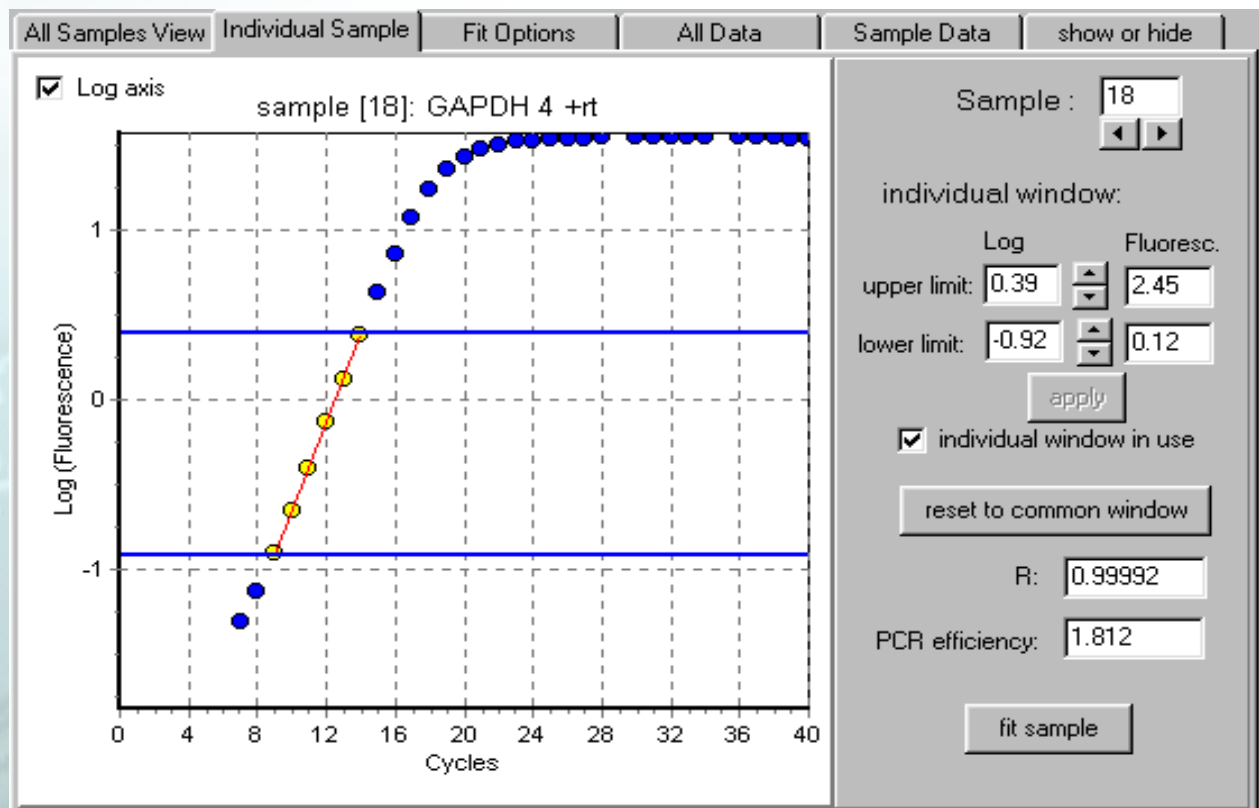
The iterative algorithm selects at least 4 and maximum 6 data points with the highest R value and a slope close to the maximum slope.

Intercept = $\text{Log} (N_0)$

i.e. $N_0 = 10^{\text{intercept}}$

Slope = $\text{Log} (E)$

i.e. $E = 10^{\text{slope}}$



Individual PCR Efficiencies obtained

Analysis of Real Time PCR data			version:7.5					
analysis date:10/1/2007								
fit option 2: n between 4 and 6 and best R								
sample	name	lower limit	upper limit	n included	No	PCR eff	R2	
1.000	w1 pr0.032062	0.070	0.500	5.000	2.86695E-08	1.633	0.999	
2.000	w2 pr0.031603	0.000	0.000	4.000	2.36015E-06	2.410	0.890	
3.000	w3 pr0.020179	0.010	0.000	9.000	0.016912114	0.812	0.742	
4.000	w4 pr-0.032059	0.010	0.120	4.000	4.81799E-15	2.353	0.999	
5.000	w5 pr0.00969	0.000	0.000	4.000	0.000353203	1.146	0.228	
6.000	w6 pr0.026444	0.010	0.190	4.000	3.21114E-13	2.335	0.997	
7.000	w7 pr0.000414	0.000	0.000	5.000	0.001092295	1.054	0.118	
8.000	w8 pr0.013775	0.000	0.000	5.000	0.000235117	1.188	0.480	
9.000	w9 pr0.076289	0.020	0.410	6.000	6.14127E-10	1.818	0.999	
10.000	w10 pr0.023188	0.010	0.470	6.000	7.534E-08	2.112	0.997	
11.000	w11 pr0.014803	0.000	0.010	5.000	0.000512664	1.186	0.733	
12.000	w12 pr0.011921	0.000	0.100	4.000	3.17236E-07	1.554	0.979	
13.000	w13 pr0.018601	0.020	0.480	6.000	1.87048E-07	1.849	1.000	
14.000	w14 pr-0.000882	0.040	0.280	4.000	5.23711E-08	1.961	0.999	
15.000	w15 pr0.393111	0.040	0.300	4.000	2.87522E-08	2.019	0.999	
16.000	w16 pr-0.084693	0.060	0.500	4.000	3.86264E-08	1.981	0.995	
17.000	w17 pr0.196027	0.020	0.210	4.000	2.64484E-09	2.203	1.000	
18.000	w18 pr-0.089781	0.140	0.560	4.000	7.65115E-06	1.565	0.999	
19.000	w19 pr0.022528	0.020	0.170	4.000	4.40541E-09	2.139	0.998	
20.000	w20 pr-0.004512	0.040	0.440	4.000	1.94535E-09	2.170	0.965	
21.000	w21 pr-0.001162	0.010	0.240	4.000	1.89643E-12	2.792	0.983	
22.000	w22 pr0.017121	0.020	0.310	5.000	3.50671E-07	1.919	0.998	
23.000	w23 pr0.0265	0.000	0.010	4.000	0.001697613	1.051	0.906	
24.000	w24 pr-0.053348	0.010	0.020	4.000	8.47769E-05	1.533	0.902	
25.000	w25 pr0.017639	0.000	0.010	4.000	0.000202257	1.242	0.789	
26.000	w26 pr0.022382	0.000	0.010	4.000	4.66835E-06	1.455	0.995	

The $2^{-\Delta\Delta C_t}$ method

$$\Delta C_t = (C_t_{\text{target}} - C_t_{\text{reference}})$$

$$2^{-\Delta\Delta C_t} = 2^{-[\text{treated (target - reference)} - \text{untreated (target - reference)}]}$$

$$X_{N,q} = \frac{X_T}{R_T} = \frac{X_0 \times (EX)_{CT,X}}{R_0 \times (ER)_{CT,R}} = \frac{\text{DNase-I treated target}}{\text{RHO treated}}$$

$$= \frac{\text{Efficiency of primer A [(10-80) DNase I conc] Ct A (10-80)}}{\text{Efficiency of RHO [(10-80) DNase I conc] Ct RHO (10-80)}}$$

= Normalized Target (Test)

$$X_{N,Cb} = \frac{\text{Untreated target}}{\text{Untreated RHO}}$$

$$= \frac{\text{Efficiency of primer A (0 DNase I conc) Ct A (0)}}{\text{Efficiency of RHO (0 DNase I conc) Ct RHO (0)}}$$

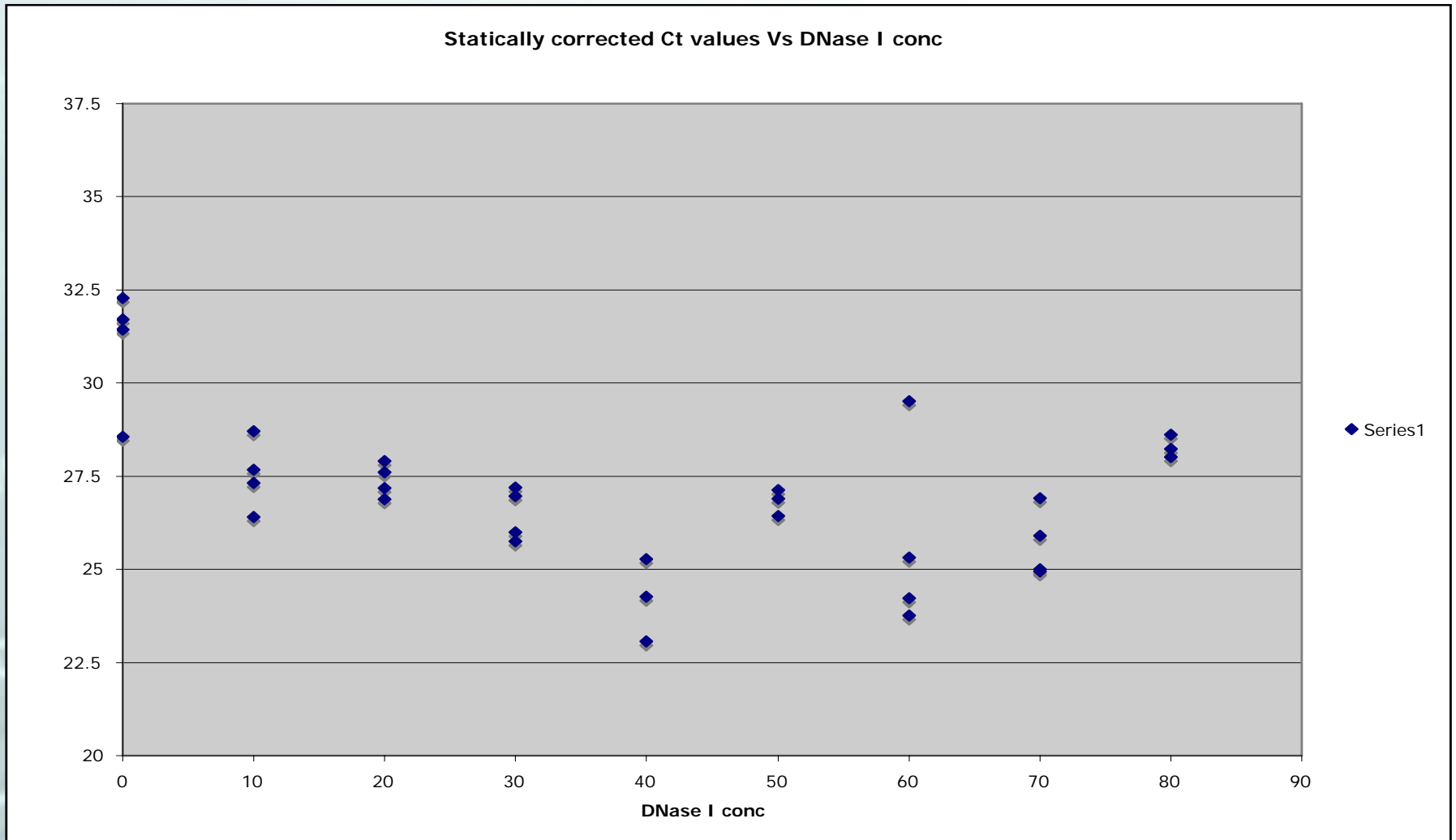
= Normalized Target (Calibrator)

$$\text{And, } \frac{X_{N,q}}{X_{N,Cb}} = \frac{\text{Normalized Target (Test)}}{\text{Normalized Target (Calibrator)}} = 2^{-\Delta\Delta C_t}$$

Data Analysis

- Efficiency corrections were performed on all test amplicons with respect to the reference amplicon.
- Efficiency corrected Ct values were then used to compute a relative copy number ratio by applying the formula
- Two primers A and B were generated across known DNase-I hypersensitive sites, which could be used as positive controls during the experiment.
- The data currently available is for the high expresser cell line MDA-MB-436 (details of which are given in the Appendix). We tested primer A within this cell line so as arrive at a single specific concentration of DNase-I which would help detect a hypersensitive site.
-
- If the same results show up on primer B, we could use this specific DNase-I concentration for detecting the Hypersensitive sites within the Kibra gene in different cell lines.

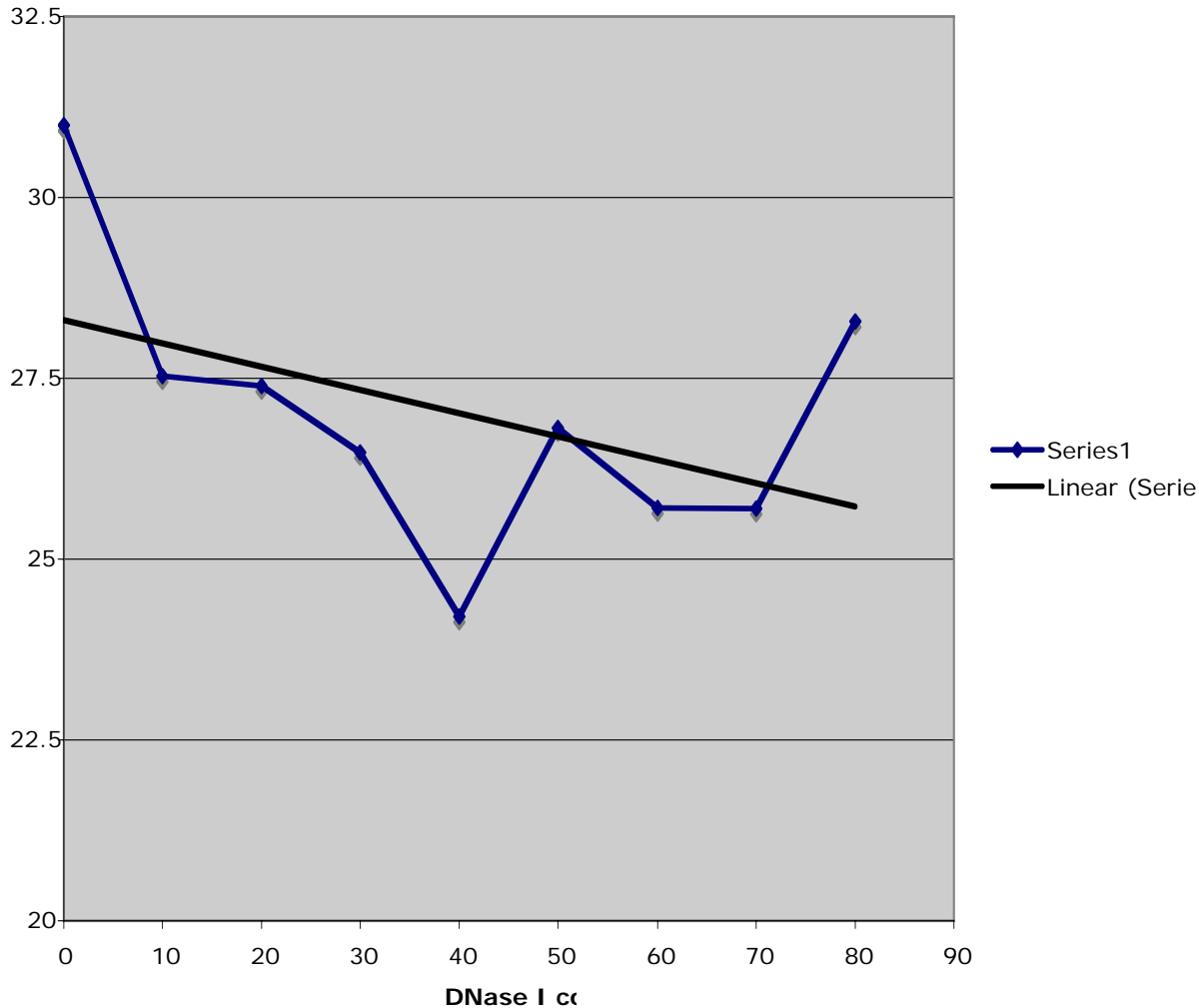
Preliminary Results



Results for Primer A (known hypersensitive region within the MDA-MB-436 (breast cancer, high expresser) cell line.

Conclusion

Avg Ct Vs DNase I



As can be seen from this graph above, there is a drop in the Ct values at DNase-I concentration of 40 U/ml. We will continue to extract more data points at this sensitive concentration so as to verify this result and also test the 2nd positive control primer B (known hypersensitive region).

Future Work - Recommendations

- We could continue to use the quantitative PCR method to elucidate the complete DNase-I hypersensitive sites within the Kibra gene, but doing so in multiple cell lines from different tissue samples, would be cumbersome.
- This is the reason why, once we arrive at a specific DNase-I concentration to be used, that would help us isolate the hypersensitive sites within the Kibra, testing this in multiple cell lines would get easier using the microarray approach. In fact up to eight different cells lines may be tested on a single microarray slide with replicates.
- The ENCODE (Encyclopedia of DNA Elements) project would also lead to providing direction in the form of developing many more high throughput approaches for detecting locus control regions within the entire human genome, thereby leading to a better understanding of the regulatory mechanism of several genes involved in disease processes.

References

- Dorschner et. al. (Dec 2004). **High throughput localization of functional elements by Quantitative Chromatin Profiling**, *Nature Methods*, Vol 1. No. 3.
- Sabo et. al. (July 2006). **Genome scale mapping of DNase-I sensitivity in vivo using tiling DNA Microarrays**, *Nature Methods*, Vol 3. No. 7.
- Livak and Schmittgen (2001) **Analysis of Relative Gene Expression Data Using Real Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method**, *Methods* 25, (Pg. 402-408).
- Jan M. Ruijter et. al. (2003) **Assumption-free analysis of quantitative real-time polymerase chain reaction data**. *Neuroscience Letters*. 339, (Pg. 62 – 66).
- Applied Biosystems (1997): **ABI User Bulletin #2 – ABI PRISM 7700 Sequence Detection System**.
- Michael W. Pfaffl **A-Z of Quantitative PCR, Chapter 3 – Quantification strategies in real-time PCR**, *International University Line (IUL), La Jolla, CA, USA*.
- Qiliang Li, Kenneth R. Peterson, Xiangdong Fang, and George Stamatoyannopoulos (November 2002) **Locus Control Regions**, *Blood*, Volume 100, Number 9, (Pg. 3077 - 3086).
- Michael W. Pfaffl (2001) **A new mathematical model for relative quantification in real-time RT-PCR**, *Nucleic Acids Research*, Vol. 29. No.9 (Pg. 2002-2007)
- Duncan Sproul, Nick Gilbert and Wendy A. Bickmore (October 2005). **The Role of Chromatin Structure in Regulating the Expression of Clustered Genes**, *Nature Reviews – Genetics*, Vol. 6 (Pg. 775-781).
- Gary Felsenfeld & Mark Groudine (January 2003) **Controlling the Double Helix**, *Nature* Vol. 421, (Pg 448 – 453).
- Gary Felsenfeld (January 1992) **Chromatin as an essential part of the transcriptional mechanism**, *Nature*, Vol. 355, (Pg. 219-224).
- David S. Gross and William T. Garrard (1988) **Nuclease Hypersensitive Sites in Chromatin**, *Annual Review of Biochemistry*, Vol. 57, (Pg. 159-197).
- Andreas Papassotiropoulos et.al. **Common Kibra Alleles Are Associated with Human Memory Performance** (2006) *Science* 314 (5798), 475.
- Suresh K. Rayala, et.al.(2006) **Essential Role of KIBRA in Co-activator Function of Dynein Light Chain 1 in Mammalian Cells**. *J. Biol. Chem.* 281: 19092-19099.

Acknowledgement

I would like to acknowledge and thank Dr. Jeffrey Touchman, for his support and guidance and for giving me the opportunity to work on this project. I am grateful to Dr. Scott Bingham and Dr. Zdzislaw Jackiewicz for agreeing to be on my Committee and providing me with their invaluable feedback.

I would like to thank the staff at Dr. Touchman's Lab in Translational Genomics (TGen) for creating an environment where students can work with other scientific officers and for trusting me to join their team. I would also like to thank my colleagues at TGen – Translational Drug Development (TD2), Karen Fisher, Heather Taylor and Ahmet Kurdoglu for sharing their experimental data with me.

Last, but not the least, I would like to thank Mrs. Renate Mittelmann, my Family and Friends....for always being there.