Transcription Factor 7-like 2 (TCF7L2) rs7903146 Polymorphism, Association with Type 2 Diabetes Mellitus Susceptibility

Naglaa Fathy Barseem1*, Mona F. El-Samalehy2 and Zeinab A. Kasemy3

1Pediatric Department, Faculty of Medicine, Menoufia University, Egypt
2GeneTic Lab. of pediatric department, Faculty of Medicine, Menoufia University, Egypt
3Department of Public Health and Community Medicine, Faculty of Medicine, Menoufia University, Egypt

*Corresponding author: Naglaa Fathy Barseem, Pediatric Department, Faculty of Medicine, Menoufia University, Egypt, Tel: 2048322170; E-mail: Naglaa_b2000@yahoo.com

Received date: February 13, 2015; Accepted date: February 24, 2015; Published date: March 2, 2015

Abstract

Objectives: The aim of this study was to investigate the possible role of TCF7L2 rs 7903146 (C/T) variant on susceptibility of T2DM among Egyptian children and adolescents.

Patients and methods: 30 T2DM pediatric patients, 20 obese children and 20 control subjects were enrolled in the study and subjected to: anthropometric measurements, routine laboratory studies including lipid profile, fasting serum insulin level and homeostatic model assessment of insulin secretion and β cell function. The rs7903146 (C/T) polymorphism was genotyped using the PCR-RFLP method.

Results: T allele of TCF7L2 rs7 903146 (C/T) was associated with T2DM in the study (P<0.001; OR=5.96, 95% CI: (2.58-16.22). Haplotype analyses showed a higher distribution of haplotype TT in the T2DM patients than the control group [56.7% vs. 15.0%, P<0.002; χ2=11.66]. Association of TCF7L2 rs7903146 and clinical and metabolic measures in T2D patients revealed significantly lower levels of fasting insulin and Homa β among carriers of T allele. Also, no significant interaction was found between T2D risk and BMI (SDS) regarding rs7903146 SNP. Conclusion: Our data prove that rs7903146 (C/T) variant of the TCF7L2 gene is associated with T2DM in our study.

Keywords: TCF7L2; Gene; Polymorphism; T2DM

Introduction

Type 2 diabetes is associated with impaired insulin secretion. Both 1st and 2nd phase insulin secretions are reduced, but the effect is particularly pronounced for the 1st phase. Although both genetic and environmental factors are thought to play a role, the processes culminating in impaired insulin secretion are not completely understood, but both genetic and environmental factors are thought to play a role. Over the past 2 years, genome-wide association scans have transformed the genetic landscape of type 2 diabetes susceptibility [1]. TCF7L2, the susceptibility gene with the strongest effect on disease susceptibility discovered to date, was conferred pre genome-wide association by Grant et al. in 2006, with rapid replication of its consequence on diabetes susceptibility in multiple populations [2]. TCF7L2 was a positional candidate gene that mapped to a region spanning 215.9 Kb on human chromosome 10q25 with replicated linkage to T2DM.

The transcription factor 7-like 2 (TCF7L2) gene is a member of the T-cell factor (TCF)/lymphoid enhancing factor family of high mobility group (HMG) box-containing transcription factors that play a key role in the Wnt signaling pathway [3]. Wnt signaling is initiated by the binding of Wnts to their receptor complex, which results in the release of catenin from its degradation complex and translocation to the nucleus. In the nucleus, catenin heterodimerizes with the TCF/lymphoid-enhancing factor family of transcription factors to regulate the expression of Wnt/beta-Catenin signaling pathway which has been implicated in blood glucose homeostasis through the regulation of pro-glucagon gene expression [4]. There is also a provided indication that active Wnt signaling stimulates the multiplication of beta cells, thus offering physiological evidence of an important role of TCF7L2 in maintaining beta cell’s mass and/or function [5]. Grant and colleagues have reported the association of (DGI05478); a common microsatellite with intrinsic region of TCF7L2 and T2DM [6]. There are at least four well-studied single nucleotide polymorphism (SNP) markers in the humanTCF7L2 gene, which are associated with T2DM, viz., rs7903146, rs7901695, rs12253372 and rs11196205. Previous studies have shown an association between T2DM and rs7903146 polymorphism of (TCF7L2) gene [6].

Abstract

Objectives: The aim of this study was to investigate the possible role of TCF7L2 rs 7903146 (C/T) variant on susceptibility of T2DM.

Patients and Methods

Study subjects

For the present study, 30 children and adolescents with T2DM, 18 males and 12 females (diabetic group) were recruited from our pediatric genetics and endocrinology outpatient clinics and those admitted to pediatric department, Menoufia University Hospitals, Egypt during the period from August 2012 to April 2014.
The study protocol had been approved by the institutional committee of ethics on human research. In our study design, 20 obese children and adolescents were enrolled to identify the role of obesity as a modulating factor for TCF7L2 T2D susceptibility (group II), 20 healthy age and sex matched children and adolescents were taken as controls including 11 males and 9 females (group III). Diagnosis of T2DM patients was based upon the American Diabetes Association criteria [7] with fasting plasma glucose ≥ 126 mg/dl, 2-h plasma glucose ≥ 200 mg/dl during an oral glucose tolerance test or plasma glucose >200 mg/dl in the presence of symptoms. All enrolled patients gave written informed consent for participation in the study.

Anthropometric assessment

All participants underwent anthropometric assessment. Anthropometric measurements were obtained using standardized techniques. Individual weight and heights were obtained from all the subjects, body mass index (BMI) was calculated using the formula Kg/m²; obesity was defined according to Cole et al. in children with BMI >95th percentile for age and sex [8]. Among obese individuals, obesity was ensured by combining data with BMI (SDS) >2. All participants were categorized according to BMI percentiles according to international survey study [8]. Waist circumference was measured to define abdominal obesity according to cut-off values for age and sex. For laboratory work up, Serum total cholesterol (TC), low density lipoprotein cholesterol (LDL-C), high density lipoprotein cholesterol (HDL-C), and triglyceride (TG) concentrations were measured by standard enzymatic methods with the use of Beckman synchron CX5 chemistry analyzer (Diamond Diagnostics, USA) [9,10]. HbA1c levels were determined by using turbidometric inhibition immunoassay. Fasting serum insulin was measured using the chemiluminescent technique (DPC kit on Immulite 1000) [11]. Homeostasis model assessment for insulin resistance (HOMA-IR) and HOMA -βindex as an approximation of insulin secretion were calculated:

\[
\text{HOMA-IR=Fasting insulin (mU/L) × Fasting glucose (mmol/L)/22.5.}
\]

\[
\text{HOMA-β}=20×\text{Fasting insulin (mU/L)/Fasting glucose (mmol/L)-3.5 [12].}
\]

Diabetes-related (islet) autoantibody testing was used to distinguish type 1 diabetes from type 2 diabetes. Type 1 diabetes is a condition characterized by a lack of insulin due to autoimmune processes that destroy the insulin-producing beta cells in the pancreas. Type 2 diabetes is primarily associated with insulin resistance. They included Islet Cell Cytoplasmic Autoantibodies (ICA) and Glutamic Acid Decarboxylase Autoantibodies (GADA) using radiobinding assays to detect antibodies to in vitro transcribed and translated antigen. With type 2 diabetes, the autoantibodies are typically absent [12].

Genotyping of TCF7L2 rs7903146 polymorphism

Sample Collection and DNA Extraction: Genomic DNA was extracted from peripheral blood from, using Bo-spin whole blood genomic DNA extraction kit "Bioflux”. DNA was librated, bounded to the Biospin membrane and then eluted after several washings [13].

**PCR-RFLP and Genotyping:** The PCR based RFLP method was employed for genotyping of the rs7903146 (C/T) polymorphism in all PCR reactions; 0.5 mg of genomic DNA was used. PCR primers used are; forward primer (5’-ATTAGAGAGCTAAGCACTTT -3’) and backward primer (5’-GAAATGTAGCAGTGAAGTG-3’).

**PCR conditions as follows:** 5 minutes at 15°C followed by 35 cycles of 45 s at 95°C, 45 s at 66°C and final extension of 20°C for 5 min. These specific PCR primers amplified 205 bp fragment in which specific restriction site determine the different alleles of rs7903146 variant. PCR products were then digested with 1 U of RsaI restriction enzyme after about 1 hr incubation at 37°C by fast digest, the resulting PCR products were visualized by 3% agarose gel electrophoresis after ethidium bromide staining. For accuracy of the genotype discrimination, some of genotyping results were validated by direct sequencing using primers which span the polymorphic region by automated sequencer (ABI, 377, Perkin Elmer) [14].

**Statistical analyses**

Mean ± SD were used to express quantitative variables. The association of rs7903146 (C/T) SNP with T2DM in the matched case-control subjects was tested by calculation of Odds ratios (ORs) with 95% confidence intervals (CIs). Mann-whitney and 1 way ANOVA were used for quantitative variables, chi-square (χ²) was used for qualitative variables. All statistical analyses were performed using SPSS statistical software (version 17.0, SPSS Inc., Chicago, IL, USA). A P value of <0.05 was considered to be significant [15].

**Results**

This study was conducted on 30, T2DM patients, 20 obese children. 20 age and sex matched subjects were taken as control. Their mean ages were 12.69 ± 2.22, 12.95 ± 1.87 and 12.75 ± 2.12 respectively. The results of our study were illustrated in the following tables and figure:

| Groups     | Test | P-value |
|------------|------|---------|
| Diabetic (N=30) | Obese (N=20) | Controls (N=20) |
| No | % | No | % | No | % |
| Sex | | | | | |

Table 1 shows the clinical and biochemical characteristics of diabetic, obese and control groups: Obese group has significantly higher BMI (SDS), waist circumference and serum cholesterol than both diabetic and control groups. Diabetic group has significantly higher fasting plasma glucose, glucose 120min, glycated haemoglobin, serum triglycerides and HDL cholesterol than both obese and control groups. Obese group has significantly higher fasting serum insulin and LDL cholesterol than diabetic group. Obese and diabetic groups have significantly higher LDL cholesterol than the control group.
Table 1: Clinical and biochemical characteristics of diabetic, obese and control groups.

Table 2 demonstrates frequency distribution of rs7903146 (IVS3C/T) polymorphism with T2D. There was a significant difference between T2D cases and obese in risk allele distribution; higher frequency of TT haplotype of 56.7 % with OR=11.49, 95% CI: (2.37-55.96) and P value=0.003, vs. 23.3 % frequency of the CT haplotype, OR=4.12 and 95%CI: 4.12(0.67-16.14). The T allele of rs7903146 variant is associated with T2DM risk (OR=5.96, 95 % CI: (2.58-16.22) and P value<0.001 in comparison to C-allele (reference allele) in obese group.

Table 2: TCF7L2 rs7903146 genetic distribution and allelic frequencies among studied groups.

Reference
Table 3 shows association of TCF7L2 rs7903146 genotype and clinical and biochemical parameters in T2D patients and controls. As compared to carriers of C- allele (individuals with CC genotype), carriers of T-allele (those with CT, TT genotype) had higher FBG and 2hr post load glucose levels, in addition to significantly lower levels of FI and HOMA-β. P-value=0.001.

|                  | Diabetic group | Controls |
|------------------|----------------|----------|
|                  | TT             | CT       | CC       | P value | TT             | CT       | CC       | P value |
|                  | Mean ± SD      | Mean ± SD| Mean ± SD|         | Mean ± SD      | Mean ± SD| Mean ± SD|         |
| Age              | 12.88 ± 2.28   | 12.71 ± 1.88 | 13.50 ± 2.66 | 0.805   | 13.25 ± 3.20   | 13.0 ± 1.82 | 12.50 ± 1.97 | 0.818   |
| BMI(SDS)         | 0.6 ± 0.48     | 0.5 ± 0.87 | 0.6 ± 0.12 | 0.459   | 0.5 ± 0.29     | 0.6 ± 0.08 | 0.5 ± 0.33 | 0.482   |
| Waist circumference (cms) | 69.17 ± 5.75 | 71.14 ± 4.87 | 67.50 ± 7.09 | 0.538   | 62.25 ± 0.95   | 71.0 ± 2.82 | 68.0 ± 5.37 | 0.038   |
| Systolic blood pressure (mmHg) | 115.59 ± 11.44 | 117.14 ± 7.55 | 112.50 ± 8.80 | 0.711   | 115.0 ± 5.77   | 113.75 ± 11.08 | 113.75 ± 8.29 | 0.966   |
| Diastolic blood pressure (mmHg) | 76.47 ± 6.79 | 73.57 ± 10.29 | 79.16 ± 8.01 | 0.455   | 77.50 ± 6.45   | 73.75 ± 7.50 | 73.33 ± 5.77 | 0.517   |
| Fasting plasma glucose (mg/dL) | 117.67 ± 10.09 | 108.41 ± 11.89 | 103.0 ± 7.07 | 0.010   | 89.0 ± 4.24    | 92.0 ± 5.34 | 85.0 ± 8.55 | 0.258   |
| Glucose 120min (mg/dL) | 201.0 ± 4.24 | 195.73 ± 25.23 | 170.50 ± 27.55 | 0.003   | 123.50 ± 10.60 | 124.0 ± 11.67 | 110.79 ± 11.12 | 0.064   |
| Glycated Haemoglobin (%) | 7.12 ± 1.16 | 7.21 ± 1.08 | 7.48 ± 1.72 | 0.839   | 4.65 ± 0.17    | 4.97 ± 0.54 | 4.70 ± 0.38 | 0.436   |
| Fasting serum insulin (IU/ml) | 8.25 ± 1.24 | 9.67 ± 0.81 | 9.90 ± 1.26 | 0.022   | 9.95 ± 0.85    | 10.40 ± 1.35 | 9.90 ± 0.84 | 0.670   |
| Serum cholesterol (mg/dl) | 157.71 ± 8.34 | 166.57 ± 10.50 | 166.50 ± 15.01 | 0.087   | 168.0 ± 17.33  | 157.75 ± 7.75 | 158.25 ± 8.61 | 0.275   |
| Serum triglycerides (mg/dl) | 132.53 ± 20.96 | 131.57 ± 18.16 | 135.67 ± 37.42 | 0.950   | 114.75 ± 22.23 | 123.0 ± 19.42 | 113.08 ± 20.14 | 0.705   |
| HDL (mg/dl)       | 38.42 ± 2.57   | 37.72 ± 2.64 | 39.06 ± 3.19 | 0.677   | 37.55 ± 3.0    | 38.37 ± 2.69 | 37.69 ± 2.17 | 0.867   |
| LDL cholesterol (mg/dl) | 118.59 ± 16.63 | 110.29 ± 16.20 | 121.17 ± 13.83 | 0.420   | 56.75 ± 11.87  | 64.0 ± 17.20 | 58.25 ± 12.21 | 0.699   |
| HOMA(IR)          | 3.64 ± 0.20    | 3.82 ± 0.36 | 4.28 ± 0.06 | 0.001   | 2.30 ± 0.38    | 2.28 ± 0.27 | 2.25 ± 0.31 | 0.959   |
| HOMAβ             | 87.92 ± 6.86   | 95.38 ± 6.77 | 98.68 ± 2.31 | 0.002   | 95.16 ± 6.34   | 109.12 ± 7.32 | 102 ± 13.76 | 0.273   |

BMI: Body mass index; HDL: High Density Lipoprotein; LDL: Low Density Lipoprotein; HOMA-IR: Homeostasis Model Assessment-Insulin Resistance; HOMAβ: Homeostasis Model Assessment beta

Table 4 shows comparison of TCF7L2 rs7903146 polymorphism in the study groups according to BMI categorization.

Regarding interaction of rs7903146 variant of TCF7L2 in all studied participants according to BMI categorization, no interaction was found between T2D risk and obesity or high BMI status regarding rs7903146 SNP of TCF7L2 gene as shown in Table 4, where obese individuals had TT haplotype frequency of 20.0 % with OR=1.32, 95% CI: (0.95-1.82), χ²=2.69 and p value=0.261. On the other hand, obese individuals showed higher significant association with C- allele (frequency of 67.5 %, χ²=3.94 and P=0.047), in comparison to lean (non over weight) individuals in our study.

Table 3: Association of clinical and biochemical characteristics of T2DM patients and control subjects based on genotype.

Table 4: Comparison of TCF7L2 rs7903146 polymorphism in study groups according to BMI categorization.

| Weight | Test | P value | OR CI 95% |
|--------|------|---------|-----------|
| Average (N=50) | Obese (N=20) |          |           |
| No | % | No | % |

Citation: Barseem NF, El-Samalehy MF, Kasemy ZA (2015) Transcription Factor 7-like 2 (TCF7L2) rs7903146 Polymorphism, Association with Type 2 Diabetes Mellitus Susceptibility. J Obes Weight Loss Ther 5: 250. doi:10.4172/2165-7904.1000250
Table 4: Comparison of TCF7L2 rs7903146 polymorphism in the study groups according to BMI categorization.

| Genotype | TT  | CT  | CC  | χ²| Ref |
|----------|-----|-----|-----|----|-----|
|          | 20  | 40.0| 11  | 22.0| 25.0|
|          | 19  | 38.0| 55.0| 20.0| 25.0|
| T        | 51  | 51.0| 32.5| 75.0| 0.047|
| C        | 49  | 49.0| 67.5| 75.0| 1.72(0.65-4.53) |

Reference

Discussion

Despite serious efforts to identify genetic variants that predispose to common forms of type 2 diabetes, till now only a few genes, such as KCNJ11, PPARG and HNF4A have been reproducibly associated with this complex disease in a variety of large-scale studies performed in the different populations [16,17]. Replication studies show that the TCF7L2 gene is an important candidate gene contributing to susceptibility to type 2 diabetes. Since the most strongly associated SNP is within an intron of the TCF7L2 gene, this SNP is not likely to be the true disease variant. However, the intronic variants may influence the risk of T2D through altering gene expression in human islets or possibly transcription regulation. TCF7L2 SNP rs7903146 is the strongest variant polymorphism associated with T2DM [18,19] and has been confirmed in numerous populations throughout the world [18,20,21]. In this study, we aimed to investigate the association between the TCF7L2 rs7903146 (C/T) variant and T2DM among study groups. Our case-control study, along with the previous reports from...
different region of the world prove that the rs7903146 (C/T) variant of the TCF7L2 gene is associated with T2DM. The T allele of TCF7L2 rs7903146 (C/T) was observed to be significantly associated with T2D risk in our studied diabetic patients (p<0.001, OR=6.74, 95% CI: (2.63-15.91).

This finding is consistent with the previous reports from other populations, except for Arab population of Persian Gulf [22,23]. This reflects the role of other genetic factors that influence the disease risk among these populations [22].

T allele frequency for rs7903146 was 32.5% in the control subjects and 68.5% in T2DM subjects of this study. The frequency of the minor T allele in the control subjects was consistent with that of rs7903146 T in the populations of Icelandic, Palestinian, Danish, Asian-Indian, Dutch, (UK-resident South Asians and Emirati populations [30,4,29,3,29,26.1,28,29 and 37.2%, respectively) [18,21,23-25] but strikingly different from that reported in the Japanese (4.2%) and Chinese (≈2%) populations [6,26]. The frequency of the minor T allele in the diabetic subjects was consistent with that of rs7903146 T among broad ethnic backgrounds, including for example populations of UK [27], Dutch [25], Amish [28], Finnish [29], Swedish [30], French [31], and US [31,32], Indian [33], and Japanese [26] origin but strikingly different from that reported in Pima Indians and Chinese diabetics [6,34]. The variability in the allele frequency, which probably reflects different ethnic backgrounds, would partly account for the discrepancies found in the results among different studies. Other factors include small sample size, age of subjects [26] and effects of environmental factors, such as lifestyle.

Haplotype analysis involving rs7903146 (IVS3C/T) of TCF7L2 in our study participants revealed a significant association of the T-allele and T2DM, compared to obese (χ²=16.07 with p value<0.001). Obese group showed no significant p-value for TT genotype.

Reinehr et al. 2004 found that reduction of overweight improves glucose metabolism in overweight children [35]. However, not all children reducing overweight demonstrated an improvement of glucose metabolism suggesting further influencing factors such as genetic markers [36]. The SNP rs7903146 in TCF7L2 is associated with an increased risk of type 2 diabetes mellitus [2,34,37-39]. The TT haplotype increases the risk of T2D in diabetic group. As well, we analyzed the association of metabolic measurements and genotype of the locus. For TCF7L2 rs7903146, mean FBG significantly increased with the number of T-alleles; CC 103.0 mg/dl, C/T 108.41, TT 117.67 (p value=0.001). In this study, our results showed that over expression HOMA-β significantly and inversely associated with T2D risk. Functional studies demonstrated that TCF7L2 specifically risk allele (T) of rs7903146 polymorphism was associated with impaired pancreatic β cell function through regulation of insulin production and secretion pathways [40]. However, the exact molecular genetic mechanisms need yet to be determined. The rs7903146 variant showed an association with T2D, but no statistically significant interactions was observed between the genotype and BMI, HDL, LDL or TG levels. Studies have reported modulating factors on the effect of TCF7L2 variant such as lipid profile, nutrition and obesity status [41]. Similarly, Hansson et al. found that allele-specific over expression of TCF7L2 in human pancreatic islets that may cause impaired insulin secretion was not altered by either insulin sensitivity or BMI [42].

Interestingly, the TCF7L2 rs7903146 variant has been associated with decreased BMI and this variant polymorphism is thought to exert a greater effect on T2D risk in lean individuals compared to obese as speculated by Van Vliet-Oostapchouk [25]. Also, Hegalson et al. [24] reported that the rs7903146 T allele, probably an ancestral and not causative variant, tags an unidentified functional variant lying outside the studied locus. It is challenging to study gene-environmental interaction effect on β cell function and genetic variant risk for T2D. Moreover, our group will conduct knowledge to confirm the association of this SNP with T2D in Egypt.

Conclusion

We have reported the association of rs7903146 variant of the TCF7L2 gene with type 2 diabetes susceptibility in our study group. Also, HOMA-β implies functional association with impaired β-cell secretory function. More insights into molecular mechanism of action of the TCF7L2 gene are recommended to clarify the pathogenesis of type 2 diabetes for prevention of the disease with attempts that could offer an early intervention health care program among at-risk children and adolescents.

References

1. Prokopenko I, McCarthy MI, Lindgren CM (2008) Type 2 diabetes: genes, new understanding. Trends Genet 24: 613-621.
2. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, et al. (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 38: 320-323.
3. Nelson WJ, Nusse R (2004) Convergence of Wnt, beta-catenin, and cadherin pathways. Science 303: 1483-1487.
4. Prunier C, Hoeveva BA, Howe PH (2004) Wnt signaling: physiology and pathology. Growth Factors 22: 141-150.
5. Ip W, Chiang YT, Jin T (2012) The involvement of the wnt signaling pathway and TCF7L2 in diabetes mellitus: The current understanding, dispute, and perspective. Cell Biosci 2: 28.
6. Chang YC, Chang TJ, Jiang YD, Kuo SS, Lee KC, et al. (2007) Association study of the genetic polymorphisms of the transcription factor 7-like 2 (TCF7L2) gene and type 2 diabetes in the Chinese population. Diabetes 56: 2631-2637.
7. American Diabetes Association (2004) Standards of medical care in diabetes. Diabetes Care 27 Suppl 1: S15-35.
8. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ 320: 1240-1243.
9. Vikram NK, Misra A, Dwivedi M, Sharma R, Pandey RM, et al. (2003) Correlations of C-reactive protein levels with anthropometric profile, percentage of body fat and lipids in healthy adolescents and young adults in urban North India. Atherosclerosis 168: 305-313.
10. Kesklin M, Kurtoglu S, Endirci M, Atabek ME, Yazici C (2005) Homeostasis model assessment is more reliable than the fasting glucose/insulin ratio and quantitative insulin sensitivity check index for assessing insulin resistance among obese children and adolescents. Pediatrics 115: 500-503.
11. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, et al. (1985) Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28: 412-419.
12. Bingley PJ, Bonifacio E, Mueller PW (2003) Diabetes Antibody Standardization Program: first assay proficiency evaluation. Diabetes 52: 1128-1136.
13. Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual Cold Spring Harbor, NY1982: Cold Spring Harbor Laboratory.
14. Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8: 175-185.
