Genetical and statistical aspects of polymerase chain reactions

Marcus Brunnert[1], Oliver Müller[2] and Wolfgang Urfer[1]

[1]Fachbereich Statistik der Universität Dortmund
[2]Arbeitsgruppe Tumorgenetik, Abteilung Strukturelle Biologie, Max-Planck-Institut für molekulare Physiologie, Dortmund

Abstract

In this paper we describe the principles of polymerase chain reaction (PCR) and its expanding use in molecular genetic research and molecular medicine. A short introduction of exemplary applications of the PCR is connected with a discussion of the lack of PCR accuracy. We give a statistical model for the PCR and discuss estimation methods in order to quantify the lack of PCR accuracy.

Key words: Polymerase chain reaction, lack of PCR accuracy, branching process, estimation of the mutation rate.

1. Introduction

The expanding use of polymerase chain reaction (PCR) in many fields of molecular genetic research and molecular medicine implicates the analysis of the accuracy of the PCR. Incorrect PCR products are useless for further applications and therefore a statistical analysis of the PCR is of great interest. The amplification of DNA by this biochemical process consists of iterative steps which form the chain reaction. Because this biochemical process produces errors the PCR is not a deterministic process. Thus the stochastic character of the PCR bases its statistical analysis using the theory of branching processes.

In this paper we focus on the stochastic modelling of the PCR considering the event of mutation of DNA while performing this DNA amplification method which leads to incorrect PCR products. The lack of accuracy of the PCR is obviously connected with the estimation of mutation rates. Besides this, the amplification rate of the PCR can explain prior information for the estimation of mutation rate. Therefore, we discuss estimators of the mutation rate and the amplification rate.

The importance of the PCR as a technique for gene analysis using gene markers is described in section 2. In section 3 the principles of PCR are described. The stochastic modelling of PCR, the estimators of the mutation rate and the amplification rate are discussed in section 4 and 5. Finally, applications of the PCR in the field of toxicokinetics are mentioned.

2. Molecular techniques for analysis of genes using genetic markers

Different types of DNA markers are used in genetic analysis. The markers segregate in Mendelian fashion and enable gene tracking studies. The resulting genetic maps can be used in the subsequent isolation of genes and in molecular genetic diagnosis.
The first markers to be developed were the sites that lead to the restriction fragment length polymorphisms (RFLPs). Investigators in maize have shown that a potentially unlimited number of RFLPs exists, which enable plant geneticists to establish genetic maps. These developments have stimulated new statistical methodologies for locating individual loci that control quantitative characters. Urfer et al. (1999) give a statistical strategy to analyze block adjusted means of genotypes from α-designs and recommend an EM-algorithm for the estimation of the genetic effect and the location of a quantitative trait locus.

Initiation and progression of a tumor is caused by alterations on the genomic levels. Genes that inhibit tumor development and thus may contribute to tumorigenesis after loss or inactivation are called tumor suppressor genes. The knowledge of a chromosomal region that is frequently lost during the development of a distinct tumor type facilitate the identification of yet unknown tumor suppressor genes. The detection of a lost chromosomal region is experimentally performed by the analysis for the loss-of heterozygosity (LOH), i.e. the loss of one out of two different allele-specific sequences. Therefore, a heterozygous sequence can serve as a marker in the detection of LOHs and for the identification of novel tumor specific genes. Such a heterozygous marker can be identified by analytical comparison of tumor DNA with normal DNA using conventional methods.
An example for genetic markers, that are potentially useful for tumor diagnosis are the microsatellite markers. These short sequence repeats show distinct and individual specific

length and copy number throughout the genome, which might change during tumor progression as a result of replication errors in the tumor cells. Thus, these markers represent useful parameters in tumor diagnosis. Microsatellite fragments can be directly amplified by special PCR methods what results in a specific microsatellite DNA patterns. Indeed, changes in the patterns of microsatellite markers have been detected in the urine of bladder cancer risk patients before the visualization of the tumors by conventional diagnostic methods.

3. The polymerase chain reaction

The polymerase chain reaction (PCR) first developed by Kary Mullis in 1985 (Saiki et al., 1985) has entered all fields of molecular biology and molecular medicine. In this chapter the principle of PCR is described, followed by a short introduction of exemplary PCR applications and an outline of reasons for lacking of the essential accuracy.

3.1 The principle

Generally, PCR is a simple *in vitro* method for the increase of the copy number of a DNA fragment (Figure 1).

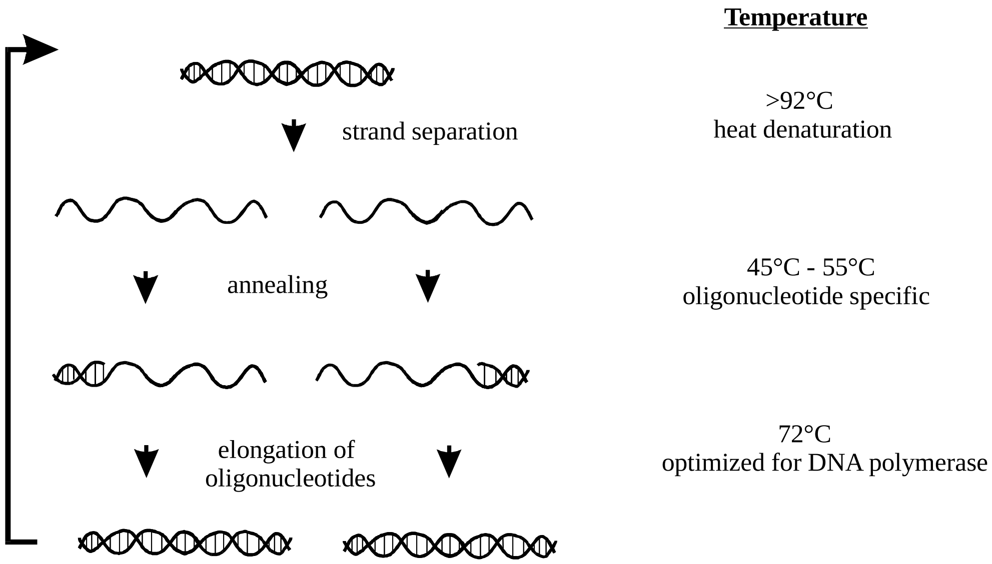

Figure 1. The principle of one PCR cycle. The strand separation is followed by the annealing of the two primers which are enzymatically elongated. The resulting products are templates in the following cycles. By repetitive cycles, the copy number is exponentially increased. The three steps differ by their different incubation temperatures.

In the first step, the so-called template DNA, i.e. the DNA fragment that has to be amplified, is denatured by increasing the incubation temperature. In this context denaturation means the separation of the DNA double strands in two single strands. In the second step, the temperature is lowered, where synthetically derived single-stranded oligonucleotides anneal to sequences up- and downstream of the DNA fragment. The sequences of these primers determine the annealing position and thus the sequence to be amplified. During the last step the enzyme DNA polymerase elongates the annealed primers by incorporation of nucleotides and synthesizing a new DNA strand that is complementary to the template sequence. These three steps form one PCR cycle by which the original copy number of the DNA template is doubled. By repetitive cycles the copy number can be exponentially increased by the factor 2^n, where n is the number of cycles. For most applications n was empirically optimized as a value between 25 and 35. In practice, the PCR process is simplified by the use of an automated heat block.

In the following short summary, the components of a PCR sample are listed. Template DNA: in diagnostic applications genomic or viral DNA; in research applications cloned genes or gene fragments. Primers: synthetic single-stranded oligonucleotides, 18 to 30 nucleotides long, which are complementary to sequences up- and down-stream of the fragment that has to be amplified. Nucleotides: dATP, dCTP, dGTP, dTTP are the DNA building blocks which are incorporated to elongate the primers and to synthesize the new DNA strand. DNA-polymerase: catalyses the DNA synthesis. To simplify the technique, a heat stable enzyme is used that is active over the repetitive cycles of increased heat incubation.

3.2 Exemplary PCR applications

The PCR technique is used in all projects when the DNA quantity is too low for analysis or for further applications. Prominent examples of PCR applications include the molecular diagnosis of tumors, the forensic person recognition and the discovery and manipulation of new genes.

- Malignant cells differ from normal cells by distinct genomic alterations which might serve as diagnostic parameters. The low amount of DNA in body fluids stemming from tumor cells prohibit the direct detection of tumor specific mutations. The increase of the copy number of the total DNA or the direct amplification of microsatellite markers by PCR allows the detection of tumor DNA in sputum, urine or feces. Thereby, tumors of the lung, the bladder or the colon can be diagnosed, respectively (Deuter und Müller, 1998). In the near future, these non-invasive techniques will at least complement the classic methods of cancer diagnosis.

- During most punishable acts, the offenders loose cells (e.g. skin, blood, hair) at the sites of crime. From these cells, DNA can be amplified and analyzed by simple PCR methods to receive an individual-specific DNA fingerprint. This fingerprint is compared with the fingerprints of suspects to identify the potential criminal. Today, the forensic area is the field with the second most PCR applications.
- All levels of molecular genetic research have been influenced by PCR. PCR applications reach from the replacement or the simplification of time consuming and error-prone gene cloning to special PCR methods which are used to alter the sequence or length of a gene.

3.3 Practical reasons for the lack of PCR accuracy

The expanding use of PCR in all fields of molecular genetic research and molecular medicine makes essential an experimental and theoretical unfolding of possible sensitivities. The most interesting aspect is the accuracy of the PCR, i.e. the consistence between the synthesized and template sequences. It is obvious that the highest possible accuracy during the DNA synthesis is necessary since low accuracy during the polymerization leads to mutations in the resulting product. A PCR product with mutations is useless for further applications and does not allow any conclusions regarding the template DNA of interest. Known experimental parameters with influence on the PCR mutation rate are: the enzyme, the buffer composition, and the temperature of primer annealing.

- The enzyme is the most important parameter that influences the PCR mutation rate. The best choice is an enzyme with proof reading activity, i.e. an enzyme that is able to correct its own sequence errors. Usually, these enzymes synthesize DNA with an error-rate of less than 1 mutation in 10^5 nucleotides compared to an error rate of 1 of 10^3 of an enzyme without proofreading activity.
- Many metal ions or organic compounds in the buffer (e.g. Mg^{2+}, DMSO), show influence on the DNA polymerase and its degree of accuracy. Thus, several authors recommend the titration of these components in order to optimize the conditions before starting a new PCR reaction.
- Last, the primer annealing position on the template DNA is crucial for specificity and accuracy of the PCR. The incubation at a temperature below the primer-specific annealing

temperature may lead to unspecific annealing and may be followed by the amplification of an unwanted DNA fragment.

4. Branching processes for modelling polymerase chain reactions

The theory of branching processes bases various stochastic models for polymerase chain reaction (Krawczak et al. 1989, Sun, 1995, and Peccoud and Jacob, 1998). In Krawczak et al. (1989) a branching process is used to analyse the impact of replication errors on the reliability of PCR. A stochastic model using the branching process in order to estimate the mutation rate of the PCR is described in Sun (1995). Peccoud and Jacob (1998) proposed an estimator for the amplification rate of the quantitative PCR using branching processes with migration. Both the quantitative PCR and the estimation of PCR mutation rates lead to the statistical analysis of PCR accuracy.

In order to define a special branching process with discrete time for stochastic modelling of PCR the following notation is necessary:

Let n denote the total number of PCR cycles and let S_i denote the number of identically copied sequences of a single-stranded DNA after i PCR cycles, i=1, ..., n. Then the original sequences can be defined as the 0^{th} generation sequences. The sequences generated directly from the original sequences are defined as 1^{st} generation sequences and inductively the sequences generated directly from the (k-1)-th generation are defined as k-th generation sequences, k=1, ... ,n. Furthermore X_k^n denotes the number of identically copied sequences of the k-the generation after n PCR cycles and X_k^n defines a random variable. A fraction λ of sequences that serves as DNA-templates for identical copies is called the efficiency of PCR. Moreover, let λ define as a constant in the whole PCR. From this follows that for all k, k=0 ,..., n , the behaviour of a certain k-th generation sequence in a certain PCR cycle does not depend on the behaviour of all other sequences in the same PCR cycle. Therefore the branching property referring to the events of *identical copy* or *no identical copy* can be assumed. Then the total number of identically copied sequences after n PCR cycles, $n \geq 1$, is given by:

$$S_n = S_0 + \sum_{k=1}^{n} X_k^n, \text{ where} \qquad (1)$$

$$X_k^{n+1} = X_k^n + \sum_{j=1}^{X_{k-1}^n} \xi_j \text{ with} \tag{2}$$

ξ a random variable that indicates the event of *identical copy* or *no identical* copy as it is defined as follows:

$$\xi = \begin{cases} 1 & \text{if one sequence is copied identically} \\ 0 & \text{otherwise} \end{cases} \qquad \text{and}$$

$$P(\xi = 1) = \lambda \text{ and } P(\xi = 0) = 1 - \lambda .$$

The sequence $S_0, S_1, S_2, \ldots.$ defines a branching process with discrete time a so-called Galton-Watson-process.

The PCR enables the assumption of the Markov property. Thereby the transition probabilities only depend on the realised event of the direct former PCR cycle. Therefore the defined Galton-Watson-process can also be interpreted as a special Marcov chain. The following figure demonstrates a realisation of the defined branching process with S_0=1.

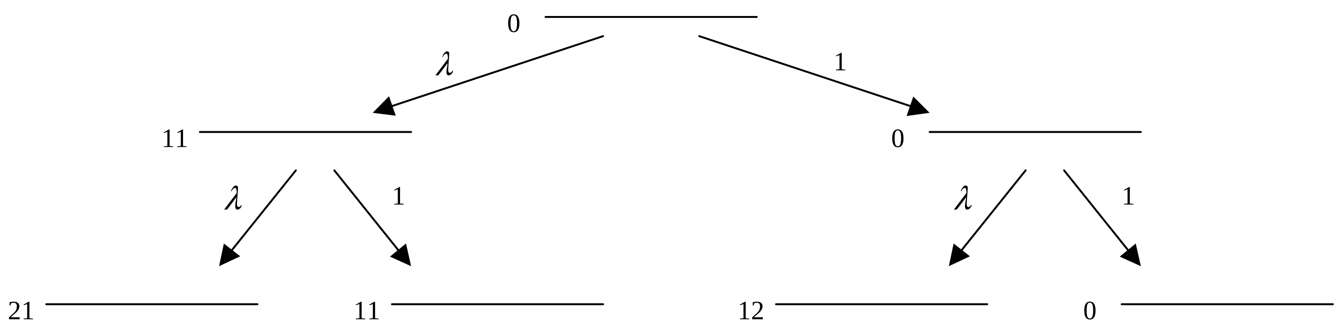

Figure 1: A branching process for two PCR cycles with one original sequence. The generated sequences and the original sequences remain with probability 1 in the PCR.

Every sequence is denoted by indexes in order to identify the sequences in the process. The first index denotes the generation of the sequence and the second index denotes the number of the copied sequences of a certain generation. If we consider the realisation in figure 1 it is

$X_0^2 = 1, X_1^2 = 2$ and $X_2^2 = 1$. This case includes all possible copy events. Another extreme case includes no copy events and yields $X_0^2 = 1, X_1^2 = 0$ and $X_2^2 = 0$.

Therefore the following results for the mean of S_n and the mean of X_k^n are important for a further analysis of the defined branching process:

$$E(S_n) = S_0 (1+\lambda)^n \tag{3}$$

and

$$E(X_k^n) = S_0 \binom{n}{k} \lambda^k . \tag{4}$$

Proof of (3):

Let $S_0=1$.

Then $P(S_1 = s_1) = \begin{cases} 0 & if\ s_1 = 0 \\ 1-\lambda & if\ s_1 = 1 \\ \lambda & if\ s_1 = 2 \\ 0 & if\ s_1 > 2 \end{cases}$.

The probability generating function of S_1 is defined as

$$E(s^{S_1}) = 0s^0 + (1-\lambda)s^1 + \lambda s^2 = (1-\lambda)s + \lambda s^2 = \varphi_{S_1}(s).$$

From this follows:

$$E(S_1) = \frac{\partial \varphi_{S_1}(s)}{\partial s}\Big|_{s=1} = \varphi_{S_1}'(1) = 1+\lambda .$$

By the property of the probability function referring to a branching process (Harris, 1963) it is:

$$E(S_n) = [E(S_1)]^n = (1+\lambda)^n .$$

Considering S_0 branching processes with an initial number of one original sequences leads to (3). ◊

The proof of (4) is described in Sun (1995).

The events *identical copy* or *no identical copy* enable a statistical analysis neglecting the knowledge of mutation. But a mutation can be one reason for the event of *no identical copy* and therefore the consideration of the distribution of mutations enables a more detailed statistical analysis.
In Sun (1995) further assumptions referring to the distribution of mutations are made:

(A1) The event of a mutation in one sequence can be assumed as randomly and rare. Then the distribution of the number of mutation can be assumed Poisson with parameter μ (mutation rate).

(A2) The length G (for example the number of bases) of a sequence influences the probability of a mutation.

(A3) The number of mutations in one sequence depends on the generation K of a sequence. For example after 2 PCR cycles there can be two mutations on the same sequence.

Assuming (A1) - (A3) the conditional probability density of the number M of mutations on a sequence is as follows:

$$P(M = m \mid K = k) = \frac{(k\mu G)^m}{m!} \exp\{-k\mu G\} \tag{5}$$

By applying the results (3) and (4) Sun (1995) showed that the distribution of the generation K of one sequence is approximately binomial with probability $\frac{\lambda}{1+\lambda}$. This is a result of strong law of large numbers. Combining (5) and the approximated (for large S_0) distribution of K Sun (1995) yield the following results for the distribution of M.

(i.) The probability density is:

$$P(M = m) = \frac{(\mu G)^m (1+\lambda e^{-\mu G})^n}{m!(1+\lambda)^n} E[Bin(n, \frac{\lambda e^{-\mu G}}{\lambda e^{-\mu G} + 1})]^m \text{, where } m \geq 0.$$

(ii.) The probability generating function of M is:

$$\varphi(s) = \frac{[1+\lambda e^{\mu G(s-1)}]^n}{[1+\lambda]^n}.$$

(iii.) E[M]=$\frac{n\lambda\mu G}{1+\lambda}$ and Var[M]=$\frac{n(\lambda\mu G)}{(1+\lambda)^2}(\mu G + 1 + \lambda)$.

(iv.) $$\lim_{n\to\infty} P\left(\frac{(1+\lambda)M - n\lambda\mu G}{\sqrt{n\lambda\mu G(\mu G + 1 + \lambda)}} \le x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-s^2/2} ds,$$

where $x \in R$.

(v.) Supposing $\lim_{n\to\infty} n\mu_n G_n = \nu$ the distribution of M is approximately Poisson with parameter $\frac{\lambda\nu}{1+\lambda}$.

The proofs of (i.) - (iii.) are described in Sun (1995). The approximations of the distribution of M are also investigated by Sun (1995). These examinations show that the normal approximation was better than the Poisson approximation for large $n\mu G$ and that the Poisson approximation almost coincided with the actual distribution after 20 or 50 PCR cycles, $\mu G = 1/40$ and $\lambda = 0.9$.

It is of great interest how far the assumptions (A1) - (A3) are recorded by the proposed result in (i.). Therefore Sun (1995) computed the probability P(M=m) as it is given in (i) for different values of G and n. These computations justify the proposed model because they show that by increasing G the probability of mutation increases, too. The same result was found for increasing number of PCR cycles. Moreover, the examinations showed that the distribution of M did not change with efficiency of PCR very much.
A supplement to the examination of Sun (1995) is the examination of the behaviour of the distribution of M referring to different mutation rates μ.

m (number of mutations)	$\mu = 9 \cdot 10^{-5}$	$\mu = 10^{-4}$	$\mu = 1.1 \cdot 10^{-4}$	$\mu = 2 \cdot 10^{-4}$
0	0.5316	0.4960	0.4629	0.2506
1	0.3319	0.3432	0.3513	0.3375
2	0.1002	0.1148	0.1289	0.2197
3	0.0097	0.0124	0.0152	0.0460
4	0.0002	0.0003	0.0004	0.0023

Table 1: The probability of the number of mutation in a randomly chosen sequence after 30 PCR cycles with a sequence length G=500 and an efficiency of PCR $\lambda = 0.9$.

Table 1 shows the change of the distribution of M with mutation rate μ. A 10%-deviation of μ causes a deviation of 10 % referring to the probability of the number of no mutation and doubling μ causes halving P(M=0). As these results have to be expected they are a further validation of the model of Sun (1995).

5. Estimation of parameters

As it is discussed in section 4 the mutation rate μ and the amplification rate of PCR are important parameters in the statistical analysis of PCR accuracy. In this section two estimators of μ proposed by Sun (1995) and one estimator of the amplification rate of PCR proposed by Peccoud and Jacob (1998) are described.

After one PCR a sample M_1,....,M_r of mutation numbers on randomly chosen sequences can be drawn. Due to the results of section 4 all M_i, i=1,...,r have identical distribution with mean $E[M_1]=\frac{n\lambda\mu G}{1+\lambda}$. By applying the moment estimation method Sun (1995) gives the following unbiased estimator of μ:

$$\hat{\mu} = \frac{(1+\lambda)\sum_{i=1}^{r} M_i}{n\lambda G r}, \qquad (6)$$

where λ, G, n follow the notation in paragraph 4.
Sun (1995) showed that the standard deviation of $\hat{\mu}$ is approximately:

$$s = \sqrt{\frac{\mu}{n\lambda Gr}(\mu G + 1 + \lambda)}.$$

Before computing this estimator two problems have to be carried. First λ is unknown and therefore λ has to be estimated. Second the mutation numbers are only available if the whole sequence is known. Carrying the second problem Sun (1995) proposed an estimator using the Hamming distance between two sequences. The Hamming distance is defined as the pairwise distance between two sequences (Sun, 1995). Both sequences are correlated through a branching process and the estimator of mutation rate is:

$$\tilde{\mu} = \frac{\sum_{i \neq j,\, i,\, j=1}^{r} H_{i,j}}{G\binom{r}{2}E(D)}, \tag{7}$$

where $E(D) = \frac{2n\lambda}{1+\lambda} - \frac{2}{(1+\lambda)S_0 + 1 - \lambda} + O\left(\frac{1}{S_0(1+\lambda)^n}\right)$.

$H_{i,j}$ denotes the Hamming distance between a sequence i and a sequence j of a sample of r sequences. Deducing of $\tilde{\mu}$ is also described in detail in Sun (1995). Nevertheless, the problem of estimation of λ is not carried.

In Peccoud and Jacob (1998) an estimator of the amplification rate m_{amp} is proposed. The amplification rate in a deterministic model of PCR is 2 if every sequence is copied identically. Then the deterministic number of PCR products after n PCR cycles is $S_n=2^nS_0$. But considering the stochastic character of PCR the amplification rate must have a distribution on the interval [1,2] which includes the theoretical cases $m_{amp}=1$ (PCR is not able to copy any sequence) and $m_{amp}=2$ (PCR is a biochemical process without producing errors).
The following estimator of the amplification rate depending on the number of PCR products of the last three PCR cycles is proposed by Peccoud and Jacob (1998):

$$\hat{m}_{amp} = \frac{S_{n-2} + S_{n-1} + S_n}{S_{n-3} + S_{n-2} + S_{n-1}}. \tag{8}$$

This estimator is deduced by Jacob and Peccoud (1996) applying a branching process with migration.

Approximately the efficiency of PCR and the amplification rate of PCR are connected by the following equation: $m_{amp} = 1 + \lambda$.

Now it is possible to quantify λ by the estimation of m_{amp}

But a simple combination of the estimations methods (6), (7) and (8) of both parameters is not possible. Measuring data in order to estimate the amplification rate requires a completley different design of experiment than measuring data in order to estimate the mutation rate and lead to complex data structures. Furthermore, the simulations in Peccoud and Jacob (1998) showed that the estimators of m_{amp} consisting of data of a few PCR cycles are not stable and especially the data of PCR's with low initial numbers S_0 enforce the use of confidence intervals for $m_{amp.}$ Moreover, taking into account the results of Peccoud and Jacob (1998) the Markov property of the described branching process in section 4 has to be discussed. This leads to further assumptions referring to λ especially the assumption of no constancy of λ.

6. Applications of polymerase chain reactions in risk assessment of potential carcinogens

The estimation of toxicokinetic parameters plays a fundamental role in the risk assessment of potential carcinogens. Urfer and Becka (1996) and Golka et al. (1999) modelled the process of chemical carcinogens into chemical active metabolites, that are able to interact with cellular macromolecules such as DNA, RNA and protein. The nonlinearity between applied dose and tumor response is supposed to be connected with the processes involved in the formation of DNA adducts. Selinski (2000) investigates the interindividual and interoccasion variabilities of toxicokinetic parameters relevant for the carcinogenicity of ethylene using an EM-algorithm.

Acrylonitrile (AN) and ethylene oxide (EO) are industrially important carcinogenic C_2-compounds whose genotoxicity is viewed in connection with their ability to bind to macromolecular targets. Both compounds are reactive towards glutathione and are detoxified via glutathione transferases (GST). Thier et al. (1999) performed a haemoglobin adduct monitoring of fifty-nine persons with industrial handling of low levels of AN. The genetic states of the polymorphic glutathione transferases GSTM1 and GSTT1 were assayed by polymerase chain reaction. A 480 bp fragment of the human GSTT1 gene was amplified with the primers 5' - TTC CTT ACT GGT CCT CAC ATC TC - 3' and 5' - TCA CCG GAT CAT GGC CAG CA - 3'. Details of the PCR reactions are given by Thier et al. (1999). The data analysis suggests that the lower EO detoxification rate in GSTT1-persons, indicated by elevated blood protein hydroxyethyl adduct levels, leads to an increased genotoxic effect of the EO background.

Further statistical methods as Kalman filter and neural network methodology used by Urfer and Schmitz (1997) and Guimaraes (1999) can be used to improve DNA sequencing accuracy. Nelson (1996) describes some efforts towards this important problem.

Acknowledgement

We are grateful to Professor Terry Speed from the Department of Statistics of the University of California at Berkely, who allowed us to use his lecture notes on 'Statistics in Genetics'. In these lecture notes we found many new ideas and problems for statisticians and geneticists.

We also would like to thank the Deutsche Forschungsgemeinschaft (SFB 475, "Reduction of complexity in multivariate data structures") for financial support.

References

Deuter, R. und Müller, O. (1998), "Detection of APC Mutations in Stool DNA of Patients With Colorectal Cancer by HD-PCR", *Human Mutation*, 11, 84-89.

Golka, K. Becka, M, Bolt, H.M. and Urfer, W. (1999), "Statistical aspects of toxicokinetics in dynamic systems: an inhalation study of propylene in rats", *Central European Journal of Occupational and Environmental Medicine*, 5, 181-191.

Guimaraes, G. (1999), "Temporal Knowledge Conversion - The Extraction of Temporal Knowledge from Multivariate Time Series ". In: *Procs of the 2nd Intl. Workshop for the Extraction of Knowledge from Databases (EKDB)*, associated with EPIA99, 65-79.

Harris, T. E. (1963), *The Theory of Branching Processes*, Springer, Berlin.

Jacob, C. and Peccoud, J. (1996), "Estimation of the offspring mean for a supercritical branching process from partial and migrating observations", *C. R. Acad. Sci. Paris Serie I*, 322, 736-768.

Krawczak, M., Reiss, J., Schmidtke, J. and Rösler, U. (1989), "Polymerase chain reaction: replication errors and reliability of gene diagnosis, *Nucleic Acids Res*., 17, 2197-2201.

Nelson, D.O. (1996), "Improving DNA-sequencing accuracy and throughput". In: *Genetic Mapping and DNA-Sequencing*, Editors: T. Speed and M.S. Waterman, The IMA Volumes in Mathematics and its Applications, Volume 81, 183-206, Springer, New York.

Peccoud, J. and Jacob, C. (1998), "Statistical Estimations of PCR Amplification Rates". In: *Gene Quantification*, F. Ferre (ed.) ,Birkhäuser, NewYork.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N. (1985), "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia", *Science*, 230(4732), 1350-1354.

Schmitz, N. and Urfer, W. (1997), "State-dependent time series models for heart rate dynamics data and their application to psychophysiology", *Informatik, Biometrie und Epidemiologie in Medizin und Biologie*, 28, 169-184.

Selinski, S. (2000), "Estimation of toxicokinetic population parameters in a four-stage hierarchical model", *Technical Report 1/2000*, University of Dortmund.

Sun, F. (1995), "The Polymerase Chain Reaction and Branching Processses", *Journal of Computational Biology*, 1(2), 63-86.

Thier, R., Lewalter, J., Kempkes, M., Selinski, S., Brüning, T. and Bolt, H.M. (1999), "Haemoglobin adducts of acrylonitrile and ethylene oxide in acylonitrile workers, dependent on polymorphisms of the glutathione S-transferases GSTT1 and GSTM1", *Arch. Toxicol.*, 73, 197-202.

Urfer, W and Becka, M, (1996), "Exploratory and model based inference in toxicokinetics". In: *Statistics in toxicology*, B.J.T. Morgan (eds.), 198-216, Oxford University Press.

Urfer, W., Mejza, S. and Hering, F. (1999), "Quantitative trait loci mapping in plant genetics by α-design experiments and molecular genetic marker systems", *Technical Report 34/1999*, University of Dortmund.