Combining DNA expression with positional information to detect functional silencing of chromosomal regions

Johannes Hüsinger1,∗, Michael Zeschnigk2, Tanja Boes1 and Karl-Heinz Jöckel1

1Institut für Medizinische Informatik, Biometrie und Epidemiologie and 2Institut für Humangenetik, Universitätsklinikum Essen, Hufelandstraße 55, D-45122 Essen, Germany

Received on September 3, 2002; revised on April 7, 2003; accepted on May 8, 2003

ABSTRACT

Motivation: Loss of chromosomal material is often observed in cancer cells. In this situation the expression of genes is related to their position on the genome. Epigenetic phenomena may also silence several genes in the same region of a chromosome. While cytogenetic or other molecular genetic methods spot changes of DNA copy number, they cannot detect other causes of silencing.

Results: We propose a method that utilizes the link from expression information gained from high-density DNA microarrays to the gene locus according to current databases. Statistical methods adequate to spot conspicuous runs of non-expressed genes are introduced and compared to one another by merit of their power and robustness against false positives.

Availability: Code for the formulae can be obtained (R code) via http://www.panix.com/~derwisch/hannes/longrun

Contact: johannes.huesing@medizin.uni-essen.de

1 INTRODUCTION

Recurrent loss of chromosomal material in a tumour is indicative for the location of tumour suppressor genes which are inactivated by structural or epigenetic alterations of the non-deleted allele. Alterations like homozygous deletion and cytosine hypermethylation can affect chromosomal regions containing several genes. Uveal melanoma is the most common intra-ocular tumour in adults. Loss of one chromosome 3 (monosomy 3), which is observed in about 50% of tumours, is strongly associated with metastatic disease of uveal melanoma (Prescher et al., 1996). Therefore, the focus of our research effort was in detecting chromosomal regions containing genes in a row which are inactivated on the non-deleted chromosome 3 and therefore perfectly suppressed.

Conventional cytogenetic and molecular genetic methods, such as comparative genomic hybridization (CGH) (Kallioniemi et al., 1992) and microsatellite analysis are used for detection of chromosomal alterations. When scanning the whole genome, deletions of down to ≈10–20 Mb length can be detected by CGH. Matrix CGH refines the resolution to below 1 Mb (Solinas-Toldo et al., 1997). Its usage for a whole genome scan is limited, however, as generating the multitude of clones is a tedious task which ranks this method still as experimental and exclusively used. Furthermore, methods that target genomic DNA fail to detect the loss of function of a gene induced by silencing on the transcriptional level, such as methylation. Functional suppression by epigenetic silencing (Paulsen and Ferguson-Smith, 2001) can affect a chromosomal region (Feinberg, 2001). In contrast, the industrial production of large-scale DNA-expression microarrays has made a shotgun approach viable. While this is not an advantage over more specific ribonucleic acid (RNA) expression detection methods when the experiment is done with specific genes in mind, it opens up new possibilities to explore cell activities at the expression level. We suggest that, with concise enough arrays, chromosomal regions associated with reduced activity can be detected.

Analysis of losses and gains of DNA material has successfully been performed using cDNA-microarrays (Pollack et al., 2002). It allows establishing a one-to-one relationship of the response of DNA and RNA to the same probe. Divergences from correlation between DNA copy alteration and RNA expression might be of interest in the aforementioned sense if they are region specific.

This work outlines a toolkit which can be used to analyse dichotomized sequenced arrangements, whose components are symbolized as 0 and 1. To that end, we used tissue from ten uveal melanoma patients with monosomy 3. The process of generating the data is described elsewhere (Tschentscher et al., 2003). The main part of the current work is a comparison of two available methods to detect regularities in the dichotomous patterns. The text starts with an introduction of the formula for the probability of finding a run of exactly
k 1s within randomly distributed \( n - m \) 0s (corresponding to expressed genes) and m 1s (non-expressed genes). A competitor to the distribution discussed here is the scan statistic which will be examined in Section 4. Section 5 further deals with the selection of a reasonable test level and the implication for the power of both methods. The application of the method in the uveal melanoma project will be investigated in Section 6. A discussion in Section 7 concludes this text.

2 PROBLEM

We are examining the regularity of occurrence of non-expressed genes with respect to their position on a chromosome or chromosomal region (in the given situation, chromosome 3). The problem of detecting a regularity is formulated by a null hypothesis of its absence. Initially, the gene expression has to be dichotomized so the chromosome is treated as a sequential arrangement of 0s and 1s. As we are interested in non-expressed genes, 1 means ‘non-expressed’ and 0 means ‘expressed’ in the rest of this paper. Due to inaccuracies in the observational setting, after dichotomizing ‘false 0s’ and ‘false 1s’ cannot be avoided. The effect of the occurrence on these false elements will be discussed in Sections 5 and 7.

Let the sample space \( \Omega = \{ (a_i)_{i \in \{1,\ldots,n\}} \in [0,1]^n, \sum_i a_i = m \} \) be the set of all possible sequences with exactly \( 0 \leq m \leq n \) 1s and \( n - m \) 0s. The null hypothesis can be formulated as follows:

\[ H_0: \text{The sampling probability is the same for each element } a \in \Omega, \text{ that is: } p(a) = \binom{n}{m}^{-1}. \]

Deviation from total randomness of the distribution of \( m \) 0s and \( n - m \) 1s can be manifold, and detection algorithms (or hypothesis tests) can perform very differently. In the current example we are interested in runs of consistently downregulated genes on a chromosomal region, which translates to an incessant run of 1s among randomly distributed 1s and 0s. The alternative hypothesis is formulated as follows:

\[ H_1: \text{There is a run of } 2 \leq k^0 \leq m \text{ consecutive 1s within the sequence. Every possible distribution of the } m - k^0 \text{ remaining 1s among the } n - m \text{ 0s occurs with the same probability } \binom{n-k^0}{m-k^0}. \]

In related problems, patterns of downregulated genes may not necessarily occur in incessant runs, but nevertheless cluster in certain regions. In this case the region can be defined as a window of width \( d \) and the alternative hypothesis can be formulated as follows:

\[ H_1^{\text{scan}}(d): \text{There is a window of width } d \text{ within which the relative frequency of 1s } k/d \text{ is systematically higher than over the whole sequence, such that } k \text{ 1s are observed within the window. The } \binom{k}{d} \text{ combinations of } k \text{ 1s and } d - k \text{ 0s in the window occur with equal probability } \binom{n-k^0}{m-k^0}. \text{ Outside the window the combination of the remaining } m - k \text{ 1s and } n - d - (m - k) \text{ 0s occurs independently of the window with constant probability } \binom{n-d}{m-k}^{-1}. \]

The null distributions of the statistics used to formulate the hypothesis tests are examined in the next two sections.

3 THE RUN STATISTIC

The run statistic, i.e. the length \( k \) of the longest sequence of 1s, is an appropriate test statistic for the hypothesis \( H_0^{\text{run}} \). It can be formalized as:

\[ k = \max_{i \in \{1,\ldots,n\}, j \in \{2,\ldots,n-i+1\}} \left[ \sum_{i=1}^{j-1} a_i \right] \left( \prod_{i=1}^{j} a_i \right). \tag{1} \]

The distribution of \( k \) can be derived by finding the number of sequences of \( n - m \) 0s and \( m \) 1s such that at least one run of length (and none longer than) \( k \) occurs. This number is a function of \( k \) with parameters \( n \) and \( m \) and will be named \( R(k|n,m) \). The discrete density of the length of the longest run \( k \) is then

\[ p(k|n,m) = R(k|n,m) \binom{n}{m}. \tag{2} \]

A formula for \( R(k|n,m) \) has been introduced (Bateman, 1948):

\[ R(k|n,m) = \sum_{t=1}^{m-k^0+1} \left[ 2f(t,k) \binom{n-m-1}{t-1} + f(t+1,k) \binom{n-m-1}{t-1} + f(t,k) \binom{n-m-1}{t} \right] \tag{3} \]

with \( f(t,k) = F(t,k) - F(t,k-1) \) where \( F \) is

\[ F(t,k) = \sum_{j=0}^{t} (-1)^j \binom{m-jk-1}{t-1} \tag{4} \]

The variable \( t \) is the number of contiguous runs of 1s in the sequence. Here, \( F(t,k) \) is the number of different sequences of \( t \) runs with a total length of \( m \) where \( k \) is the maximum length of a run.

Bateman derives the power function for the run statistic when the alternative is autocorrelation. He concludes that in the case of autocorrelation \( \text{i.e. } P(a_1 = 1|a_{t-1} = 1) \neq P(a_1 = 1|a_{t-1} = 0) \), the run test (based on \( t \), the number of runs) is more powerful than a test based on the run statistic.
In the presence of a fixed run with length $k^0$, however, the run statistic is most sensitive to detect such a run. This case, peculiar to the situation at hand, happens much less often in everyday applications than autocorrelation, therefore the distribution of $k$ is less known than the distribution of $t$, which is a textbook example.

4 THE SCAN STATISTIC

The scan statistic moves a window of pre-defined length $d$ over the whole sequence (i.e. the chromosome) and returns the maximum number of 1s $k$ over all possible positions of the window. It can be written as

$$k = \max_{j \in \{1, \ldots, n - d + 1\}} \sum_{l = j}^{j + d - 1} a_l. \quad (5)$$

Scan statistics (Glaz and Naus, 2001) have been developed for both a discrete (clusters of 1s between 0s) and continuous (events in continuous time) support. The discrete case is interesting here. The distribution of a scan statistic, assumed that $H_0$ is true, is not known generally. Naus (Naus and Wallenstein, 1974) developed a formula which holds exactly if $k \geq m/2$ and $n/d$ an integer:

$$p_{\text{scan}}(k|d, n, m) = \left(\frac{nk}{d} - m - 1\right) \mathcal{H}(k, m, d, n) + 2 \sum_{s = k}^{m} \mathcal{H}(s, m, d, n) \quad (6)$$

where $\mathcal{H}(k, m, d, n)$ is the hypergeometric distribution for the probability of finding exactly $k$ 1s in a sample of $d$ out of a population with $m$ 1s and $n - m$ 0s.

Note that, if $k \geq m/2$ holds, the formula for $p_{\text{scan}}(k|k, n, m)$ will give the same result as the run statistic $p_{\text{run}}(k|n, m)$ (expression 2). Under this condition no cluster of more than $k$ 1s outside the window can exist. For $k < m/2$, the formula neglects the case that several clusters of size $k$ or more exist in different positions of the window. The formula counts these sequences as multiple instances. Therefore, in the case $k < m/2$, the general condition for a probability function $\sum_k p_{\text{scan}}(k|d, n, m) = 1$ is violated. It is claimed (Naus and Wallenstein, 1974) that for $p_{\text{scan}}(k|d, n, m) < 0.1$ the formula still gives a good approximation (as the probability of multiple clusters is low if the relative frequency of 1s is low).

As the distribution naturally depends on the window width $d$, statistics have been proposed to measure the deviation from the null hypothesis of no clustering over all window sizes. However, a probability distribution of these generalized scan statistics could not yet be established. Thus, distribution functions and $p$-values have been derived from simulations (Nagarwala, 1996).

In the application at hand, however, a reasonable window size can be inferred. It depends on the resolution of the chip in terms of subsequent genes and on the resolution in base pairs of the concurrent CGH (if a CGH is performed). In the current application, with about 400 genes on the chip whose locus on chromosome 3 could be extracted (see Section 6), a chromosome of about 200 Mb length and a resolution of about 10 Mb of standard CGH, a window width of 20 looks feasible.

Also, the window has to be chosen wide enough so that the probability is low so that it will be filled with 1s. A window of width 5 is clearly too narrow when $n = 100$ and $m = 40$.

5 SELECTION OF CRITICAL VALUES AND POWER CONSIDERATIONS

The derivation of the distribution function for the run and scan statistic under $H_0$ allows us to derive precise likelihoods of run or scan statistics for predefined chromosomal regions with a fixed number $n$ of genes probed by the chip. Any use of a detection algorithm has to cope with the occurrence of false positives or negatives, and to pick a detection threshold accordingly. The positive predictive value (i.e. the probability of dealing with an actual case of silencing, given the threshold for the test statistic $k$ has been exceeded) depends on both the likelihood (i.e. the sum of the probabilities from $k$ to infinity, assumed $H_0$ is true) and the prior probability that a silencing event occurs.

The assessment of prior probabilities is currently based on very little information. Estimations on the frequency of methylation in the promoter region of genes differ wildly. While the rough assessment of magnitude of methylation events is in the ‘hundreds’ (Garber, 2002) in the human genome, we assume to expect every 25th to every 1000th gene to be affected. When a posterior probability of 0.95 is desirable that a run of non-expressed genes can be attributed to methylation, the likelihood of a run of non-expressed genes occurring by chance has to be set to between $2.5 \times 10^{-3}$ and $2 \times 10^{-3}$ to achieve a posterior of 0.95 that a run of genes is suppressed by methylation. As methylation is known to happen in CpG islands (in mammals), the likelihood can be set higher for long runs of unexpressed genes known to contain CpG islands. Other possible sources of silencing, such as deletion of small regions, are to be taken into account so that one would like to increase the prior probability here, and thus lower the threshold.

Next, the distribution of the run statistic (expression 1) is compared to the scan statistic (5) with respect to power. Based on the considerations given above for a critical value $\alpha$, one can pick a critical value $k_{\text{run}}(\alpha)$ and $k_{\text{scan}}(\alpha)$ (longest run) for $k$ using the cumulative distribution function from the formulae.
for the discrete densities (2) and (6):
\[
\begin{align*}
  k_{\text{scan}}^* (\alpha) & = \min\{k : 1 - P_{\text{scan}}(k-1|d,n,m) > 1 - \alpha\} \\
  k_{\text{run}}^* (\alpha) & = \min\{k : 1 - P_{\text{run}}(k-1|n,m) > 1 - \alpha\}
\end{align*}
\]

(7) \quad (8)

We are deriving the power under the assumption that \( H_1^* \) is true. This criterion clearly favours the run statistic over the scan statistic. We perform this comparison because it is most important to us as the alternative describes the situation we want to investigate. Even if one is interested in detecting the alternative \( H_1^* \), using the longest run as a test statistic might not be the wisest choice: the scan statistic is advantageous in that it is less sensitive to false up-regulations generated by cross-hybridization or measurement error. For the application it is of interest how much power is lost when the more robust scan statistic is chosen.

The length \( k^0 \) of the ‘true run’ of silenced genes can be overestimated by an incidental downregulation of neighbouring genes. The power of the tests, given that \( k \geq k_{\text{scan}}^*(\alpha) \) (\( k \geq k_{\text{run}}^*(\alpha) \)), is of course 100%. For true \( k \) less than the critical values, the probability of having enough 1s in the neighbourhood of the true run has to be determined to calculate the power. Effectively, we are searching for the conditional probabilities
\[
P_{\text{run}}(K \geq k_{\text{run}}^*(\alpha)|K \geq k^0,n,m)
\]
and
\[
P_{\text{scan}}(K \geq k_{\text{scan}}^*(\alpha)|K \geq k^0,n,m,d)
\]

Both formulae are derived using the assumption that the probability of a run of length \( k \) (\( k \) being generic for \( k_{\text{run}}^* \) and \( k_{\text{scan}}^* \)) not containing the true run is negligible. This assumption is equivalent to the prerequisite of formula (6) being a good approximation even if \( k < m/2 \).

First, the conditional probability for the longest-run statistic
\[
P_{\text{run}}(K \geq k_{\text{run}}^*(\alpha)|K \geq k^0,n,m) - \sum_{k=k^0}^{m} P_{\text{run}}(k|n,m) \sum_{k=k^0}^{m} P_{\text{run}}(k|n,m)
\]

(9)

where the denominator \( \binom{m}{k} \) has been cancelled. Further simplification of the ratio of these expression does not seem possible as this would require factorizing expression 3. This makes it look unlikely that computation time of this expression can be coerced below \( O(n^3) \).

Now the distribution function of \( K \) in the scan statistic setting is derived. It is helpful to consider all windows of length \( d \) that contain the \( k^0 \) 1s (see example in Fig. 1). The union of these windows spans a sequence of length \( 2d - k^0 \), where \( 2d - 2k^0 \) elements are outside the true run. Let the number of 1s within these \( 2d - 2k^0 \) elements be \( m^* \). Then, we are looking for the probability that the maximum number of 1s in the window outside the true run is at least \( k^0 \) out of \( m^* \). This leads to the distribution for the scan statistic conditional on \( m^* \), which itself follows a hypergeometric distribution. The resulting discrete density is therefore:
\[
P_{\text{scan}}(K \geq k_{\text{scan}}^*(\alpha)|K \geq k^0,n,m) = \sum_{k=k^0}^{m} \sum_{m^*=0}^{m-k^0} P_{\text{scan}}(k|d-k^0,2d-2k,m^*) \times H(m^*|n,m,2d-2k)
\]

(10)

Note that within the above formula, the probability function \( P_{\text{scan}}(\cdot) \) may not be replaced by the right hand-side of (6) in all cases as the prerequisites are violated. The total number of genes within the subsequence is so low, however, that results can be obtained by counting the number of cases for denominator and numerator.

To compare the power of the scan statistic and the run statistic, the argument of the distribution function has to be set to the critical value, which is denoted \( k^* \) in Table 1. Power curves are given in Figures 2 and 3. With \( m \) selected as being \( n/4 \),

**Fig. 1.** An example of a power calculation using a true run of four 1s in a sequence of \( n - m = 12 \) 0s and \( m = 12 \) 1s and a window width of 8. The number of arrangements under which at least \( k_2 \) 1s are found within the window can be found by taking out the block of four 1s and scanning over the four positions on the left and right of the block to detect at least \( k_2 - k^0 \) 1s.

**Table 1.** Critical values in \( m \) out of \( n \) 1s for \( \alpha = 0.05 \): run statistic versus scan statistic

<table>
<thead>
<tr>
<th>( n )</th>
<th>40</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m,d )</td>
<td>10,10</td>
<td>20,20</td>
<td>30,30</td>
</tr>
<tr>
<td>( k_{\text{run}}^* )</td>
<td>5</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>( k_{\text{scan}}^* )</td>
<td>7</td>
<td>15</td>
<td>26</td>
</tr>
</tbody>
</table>

2d - k^0
Fig. 2. Power of the run statistic. The length of a true run of 1s is put into the x-axis, the y-axis gives the probability that stray 1s adjacent to the true run extend it to a run long enough to reject the null hypothesis. The power curves tail off at $k^0 = k_*$, above which the power is automatically 100%.

Fig. 3. Power of the scan statistic. The length of a true run of 1s is put into the x-axis, the y-axis gives the probability that stray 1s adjacent to the true run extend it to a run long enough to reject the null hypothesis. The power curves tail off at $k^0 = k_*$, above which the power is automatically 100%.
the power cannot be greater than 25% when \( k_0 \) is by one less than the critical value \( k^* \). A satisfactory power with \( k_0 < k^* \) can only be achieved when \( m/n \) is high.

## 6 IMPLEMENTATION

In the investigation of monosomy 3 in uveal melanoma cells, we looked at the expression patterns of the genes on chromosome 3. We used Affymetrix U95A version 2 chips. The average difference as a summary measure generated in Micro Array Suite 4.0 by Affymetrix was taken for the expression. A gene was marked as non-expressed in a sample if the maximum of expression over all probe sets (for the case of multiple probe sets per gene) was below 500. The distribution of the ‘Average Difference’ measure calculated by matches and mismatches of non-expressed genes is theoretically symmetrically distributed around 0. The threshold of 500 was estimated to be roughly at the 97th percentile of non-expressed genes. This corresponds to a probability of about two-thirds that a true run of non-expressed genes will not get overlooked.

Information about the location of the genes was taken from the Bioconductor suite (Gentleman and Carey, 2002) using the annotation library HG-U95A, Version 1.1.1, which is built upon UniGene Build 158 ftp://ftp.ncbi.nih.gov/repository/UniGene/Hs.data.gz and the LocusLink database ftp://ftp.ncbi.nih.gov/refseq/LocusLink/LLtmpl.gz dating from January 30, 2003. All genes with a probe set representation on the HG-U95A chip which have a chromosome location entry were used. The total number is 407 genes. Of these, between 191 and 216 genes fulfilled the criterion of being unexpressed.

With roughly 50% of the genes expressed and a length of 11–13 as an \( \alpha = 0.05 \) critical value for the run statistic, the method currently has an average resolution of about 6 Mb, depending on the density of the genes used in the analysis (see Fig. 4 and Table 2).

Five out of ten tumours had the longest run of non-expressed genes around 60 million base pairs (Mb) which indicates a number of systematically downregulated genes in that region. The run lengths, however, do not give enough support to refute the hypothesis that the lack of expression is not related to the...
fact that these genes are next to each other on the chromosome. Moreover, most of the genes in these regions were also downregulated in the collective of tumours lacking the monosomy 3 aberration.

7 DISCUSSION

The problem of investigating the non-randomness of runs is common in sequence analysis and has substantially initiated methodological research in combinatorics and statistics. The scan statistic, while being a relatively old concept (Krauth, 1997) has also been revisited in the context of both the nucleotide and gene (Hoh and Ott, 2000) level. The run statistic is much less often used in the genetic context and elsewhere.

The null hypothesis of unrelatedness of position and function is known not to hold strictly. Instead, some regions exist with consecutive genes that are functionally related. These regions could account for false positive findings. Our experience is that these incidences of false positives is rather rare, and moreover are interesting in their own right. Recently, a domain-dependent expression pattern of genes has been observed (Caron et al., 2002). There, the authors took a continuous rather than a dichotomized measure to detect this pattern. For want of a closed-form distribution assumption, they used a randomization test argument to assess the amount of non-randomness.

False negatives can happen in the case of functional silencing when the array detects an upregulation of a gene whose expression is in fact suppressed. False positives because of lack of detection of an expression are much less detrimental to the size and power of tests and therefore for the statistical significance of findings if the lack of detection is independent of the position. This is a rather safe assumption because the detection of existing RNA is unrelated to the gene locus. If one is concerned about false negatives, i.e. detection of an expression where none is present, one will be better off using the more forgiving scan statistic which indicates clustering rather than incessant runs.

The range of applications of this approach is quite narrow, as deletions visible in a karyogram are not often scanned for invisible causes of inactivation on the corresponding chromosome. One could think about extending the applicability to detecting regions of inactivation of one of a pair of chromosomes: runs of lower-than-average expressed neighbouring genes on a chromosome can be used to detect local loss of heterozygosity. To that end, the threshold for expression intensity would have to be set to a higher value, such as the mean of expression levels of control tissue samples. In this setting, carefully picked control samples could allow this technique to be applied in cDNA settings, looking for ‘red regions’ or ‘green regions’. As the impact of homozygous deletion or multiplication on expression is (in contrast to biallelic silencing) not deterministic but only causes corresponding expression activity on the average, the longest-run statistic will have to be replaced by the scan statistic to be applicable. The amount of correlation between DNA and RNA is investigated in Pollack et al. (2002).

When scanning the whole genome for conspicuous runs, it makes arguably sense not to condition on different relative frequencies of unexpressed genes for each chromosome, if one would not assume the probability of a gene being expressed to vary with the chromosome. It seems natural then to assume a binomial distribution with the probability set to the relative frequency over all chromosomes. For binomial runs, a distribution of the length of the longest run k can be derived from a formula given by Godbole (1991):

\[
P_G(k|n, p) = 1 - \sum_{y=\lceil n/k \rceil}^{n} (1 - p)^y p^{n-y} \times \sum_{j=0}^{(n-y)/k} (-1)^j \binom{y + 1}{j} \binom{n - jk}{y}
\]

The unconditional distribution has asymptotic behaviour which is well described via the Poisson approximation to the binomial. The adequacy of a common null distribution for different chromosome can be disputed, however, in the light of recent results (Caron et al., 2002) showing chromosome-specific gene densities and expression intensities.

A similar approach has been suggested that looks for consistently misregulated genes across different tissue samples (Wahde et al., 2002). If the suppression mechanism occurs regularly in conjunction with the disorder under investigation, the k consecutive genes found in a specific sample can be scanned for consistency in more samples using the results mentioned there. Aside from novel regions detected by the algorithm proposed here, one also can feed results from databases of cytogenetic aberration [such as Progenetix (Baudis and Cleary, 2001)] into the approach in Wahde et al. (2002).

ACKNOWLEDGEMENTS

This work was supported by the Federal Ministry for Education and Research through the NGFN project ‘Expression Profiling of Human Cancer by DNA-array Analysis. Frank Tschentscher performed the experiments and kindly provided the data. We are grateful to three anonymous referees, one of them especially helped a lot to eliminate mistakes and improve the manuscript by thorough treatment of preliminary versions.

REFERENCES


