IL NUOVO CIMENTO Vol. 35 C, N. 5 Suppl. 1 Settembre-Ottobre 2012
DOI 10.1393/ncc/i2012-11339-9

Colloquia: PR PS BB 2011

A variational piecewise smooth model for identification of chromosomal imbalances in cancer

S. Morganella and M. Ceccarelli

Department of Biological and Environmental Studies, University of Sannio - Benevento, Italy
Bioinformatics CORE, BIOGEM s.c.a.r.l. - Ariano Irpino, Italy

ricevuto il 30 Settembre 2011; approvato l' 1 Dicembre 2011

Summary. — Monitoring of changes at the DNA level enables the characterization of the underlying structure of genetic diseases. In particular, copy number alterations (CNAs) are increasingly being recognized as an important component of genetic variations in cancer: oncogenes may be enhanced by DNA amplification and tumor suppressor genes may be inactivated by physical deletion. Encouraged by the advent of array comparative genomic hybridization technology, several biological studies have been designed to look for chromosomal aberrations involved in cancer. Hence, the development of algorithms aimed at the identification of CNAs is a current challenge in bioinformatics. Despite the amount of proposed approaches, identification of CNAs is yet an open problem. Here we propose a new approach for detection of CNAs that extends a previously published algorithm where a popular image segmentation variational model was used. The proposed algorithm, called Vega Multi-Channel (VegaMC), starts from the assumption that copy number profiles are piecewise constant and finds the optimal segmentation by minimizing a functional energy that represents a compromise between accuracy and parsimony of the boundaries. We applied VegaMC on a published gastrointestinal stromal tumor aCGH dataset, showing the ability of the proposed approach in the identification of well-known cytogenetic mutations, and eventually discover new ones.

PACS 87.18.Nq – Large-scale biological processes and integrative biophysics.
PACS 87.18.Vf – Systems biology.
PACS 87.85.Ng – Biological signal processing.

1. – Introduction

Chromosomal abnormalities are known to have a direct influence on the development and progression of genetic diseases as cancer [1]. In the latest years, greater attention has been paid to CNAs which are defined as genomic regions, larger than 1 kb, in which copy number differences are observed between two or more genomes [2]. Indeed, it was observed that oncogenes are often located in regions with a gain in their copy number, in contrast, oncosuppressor genes are found in lost chromosomal regions. In addition, copy-neutral loss of heterozygosity (LOH) is receiving greater attention as a mechanism of possible tumor initiation [3]. Advent of array comparative genomic hybridization (aCGH) technology has allowed the monitoring of changes at the DNA level for more than one million of chromosomal loci (probes). In particular, aCGH provides an indirect

TABLE I. – *List of JS-algorithms.*

Algorithm	Preprocessing	Output	Ref.
BSA	No	Statistical evidence (q-value)	[15]
chgMCR	Yes	None	[9]
CNAnova	No	Statistical evidence (p-value)	[18]
GADA	No	None	[16]
GAIA	Yes	Statistical evidence (q-value)	[12]
GISTIC	Yes	Statistical evidence (q-value)	[10]
JISTIC	Yes	Statistical evidence (q-value)	[11]
KCsmart	No	Statistical evidence (q-value)	[17]
pREC-A	Yes	Probability of alteration	[13]
pREC-S	Yes	Probability of alteration	[13]
VegaMC	No	Weighted mean	proposed

measure of copy number for each probe, this measure is known as Log R Ratio (LRR) and is computed by the ratio of observed to expected hybridization intensities. Supported by effectiveness of aCGH, several biological studies aimed at the identification of functional CNAs in cancer have been proposed [4-6]. These studies produced a vast amount of data that need interpretation, hence a current challenge in bioinformatics is the development of algorithms to analyze these data. In the literature, we find algorithms that are designed to provide a map of CNAs of a single genome [7,8] (the so-called segmentation algorithms) and algorithms that, as VegaMC, look for CNAs shared among the cohort of available subjects. In this work we focus our attention on this last kind of algorithms that jointly analyze copy number data from multiple samples. For brevity, we will refer to these algorithms as Joint Segmentation algorithms (JS-algorithms).

Among the proposed JS-algorithms, we find methods that use a preprocessing step to obtain segmented input data and methods that directly work on the original LRRs. Table I reports a summary of some JS-algorithms for which a software implementation is available. In cghMCR [9] a preprocessing step is used to obtain smoothed LRRs. Smoothed data are then used to distinguish between normal and altered probes by a percentile-based approach. Finally, cghMCR returns the regions found aberrant at least in 75% of the samples. GISTIC [10] and JISTIC [11] also need smoothed LRRs: they use smoothed data to compute for each probe a statistic representing the strength of the aberration. This statistic is then used in a conservative permutation test to assess the statistical significance of regions and finally CNAs that have a high evidence to be sites of driver mutations are extracted by an iterative procedure known as peel-off that represents the main difference between GISTIC and JISTIC. GAIA [12] is very similar in spirit to GISTIC and JISTIC, indeed, it also uses a conservative permutation test, but GAIA uses as input a discrete representation of the observed LRRs, so that it can distinguish between normal and aberrant probes. GAIA exploits this discrete data representation to obtain information on the degree of homogeneity between adjacent probes, so that, peel-off procedure can take in account both statistical significance and within-sample homogeneity. Also pREC-A and pREC-S [13] use a preprocessing step but, differently to other approaches, this step is used to compute a probability of alteration associated to each probe and not a smoothed or discretized representation of the LRRs. These probabilities are used in a Hidden Markov Model (HMM) to compute the joint probabil-

ities of alteration for adjacent probes, in particular pREC-A and pREC-S use the HMM to estimate the probability that consecutive probes belong to the same aberrant region.

The use of segmented data has been criticized because of the potentially loss of information that it entails when input data are strongly affected by noise [14] and, in addition, there is no possibility to avoid the propagation of the error from the preprocessing to the rest of the downstream analysis [13]. To overcome these problems JS-algorithms, directly working on the original LRRs, have been proposed. In BSA [15], segmentation and copy number assignment are simultaneously performed by using an hierarchical Bayesian model. BSA sequentially detects CNAs based on the assumption that observations for different subjects in different genomic positions are independent conditional on the boundaries. Also GADA [16] works on the original LRRs. GADA performs a decomposition of the observed LRR into three components: the change in hybridization due to altered copy number, the reference hybridization intensity for non-aberrant probes and the noise component modeled as a zero-mean Gaussian process. GADA is based on the assumption that copy number hybridization component can be approximated as a piecewise constant (PWC) function and it uses an expectation maximization framework to jointly estimate all three components. KCsmart [17] separates the original LRRs in positive and negative values so that they can be separately summarized across the samples. Finally a smoothed estimation of CNAs is obtained by using a flat top Gaussian kernel function. A different approach was proposed in CNAnova [18], indeed, it assumes that the dataset is composed by a set of cancer samples and by a set of normal reference samples. Based on this assumption and starting from the observed LRRs it computes the distributions of the F- and t-statistics by using the one-way analysis of variance (one-way ANOVA). These distributions are used to detect CNAs.

Here we describe a new JS-algorithm (VegaMC) that does not need of a preprocessing step and jointly analyses copy number data. The proposed approach extends a previously published segmentation algorithm (Vega [8]) so that more than one sample can be analyzed and recurrent CNAs can be detected. VegaMC is based on an efficient image segmentation model known as Mumford and Shah model [19] where the PWC assumption is used to define a functional where both accuracy and parsimony of the boundaries are considered. Segmentation is performed by minimizing this functional with a greedy region growing approach where the sequence of the regularization parameters is identified by a data-drive heuristics. The difference between VegaMC and Vega is on the energy functional: in Vega the functional is defined on a scalar function, while in VegaMC this functional is defined on a vector-valued function whose components are different channels (each channel models an aCGH sample). We applied VegaMC on a recently published aCGH dataset of gastrointestinal stromal tumor composed by 25 samples. Current knowledge about this tumor has been used to validate the performance of VegaMC showing the ability of the proposed approach in detection of well-known aberrant cytobands.

The paper is organized as follows, the next section reports the datails of the adopted method and the corresponding minimization algorithm, whereas the section of results reports the obtained segmentation on real data.

2. – Method

This section is structured as follows: first we describe the Mumford and Shah model which represents the basis of the proposed approach, then we provide a detailed description of VegaMC.

2˙1. *Mumford and Shah Model.* – In their original work, Mumford and Shah [19] introduced the basic properties of variational models for computer vision. The aim of these models was to provide a set of mathematical foundations for appropriate decomposition of the domain Ω of a function $u_o : \Omega \to \mathbb{R}^n$ (*i.e.* an multichannel image) into a set of disjoint connected components Ω_i:

$$\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \Omega_n,$$

such that the function u_o varies smoothly within each connected component Ω_i and u_o varies discontinuously across the boundary Γ between different Ω_i. Therefore, the decomposition (segmentation) problem may be viewed as seeking of mathematical models to define and compute optimal approximations of u_o by piecewise smooth functions. To this purpose, Mumford and Shah proposed the minimization of the following functional:

$$E(u, \Gamma) = \alpha \int_{\Omega} (u - u_o)^2 \mathrm{d}x \ \mathrm{d}y + \int_{\Omega \setminus \Gamma} | \bigtriangledown u|^2 \mathrm{d}x \ \mathrm{d}y + \lambda |\Gamma|, \tag{1}$$

where α and λ are two non-negative parameters weighting the different terms in the energy: the first term requires that u approximates u_o, the second term takes in account the variability of u within each connected component Ω_i and the third term penalizes complex solutions in terms of the length of the boundaries $|\Gamma|$. A special case of eq. (1) is obtained when the approximation u of the signal u_o is considered to be a piecewise constant (PWC) function (u constant within each connected components Ω_i). For this case Mumford and Shah proposed the so-called *piecewise constant Mumford-Shah model*:

$$E(u, \Gamma) = \sum_i \int_{\Omega_i} (u_o - u_i)^2 \mathrm{d}x \ \mathrm{d}y + \lambda |\Gamma|. \tag{2}$$

It is easy to prove that the minimum for this model can be obtained by posing u_i as the mean of u_o within of each connected component Ω_i. Analysing the functional (2), we can observe that the first term takes in account the error in the approximation of u_o, while the second term penalizes complex segmentations, thence this functional represents a compromise between the accuracy of the approximation and the parsimony of the boundaries. It is important to notice that the resulting segmentation depends on the scale parameter λ, indeed it determines the amount of computed regions: when λ is small many boundaries are allowed so the resulting segmentation will be *fine*, while as λ increases the segmentation will be coarser and coarser.

2˙2. *VEGA Multi-Channel (VegaMC).* – For a better comprehension of the aim of the proposed approach consider the situation depicted in fig. 1 which shows the LRRs of 4 samples. In this dataset we can distinguish 5 regions where, in agreement with the Mumford and Shah model, in each region the LRRs (the function u_o) varies smoothly within each of them and varies discontinuosly across the boundaries. For non aberrant probes LRR = 0 is expected (red line) and deviations from this value are indicator of possible mutations. In the example depicted in fig. 1 we found 3 non aberrant regions (R_1, R_3 and R_5) and two possible aberrant regions (R_2 and R_4). Aim of VegaMC is the identification of these regions by a joint analysis of all available samples.

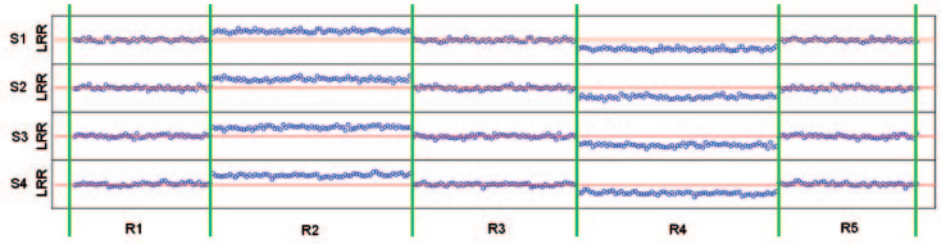

Fig. 1. – Example of aCGH dataset composed by 4 samples. The red line indicates the value 0, which represents the LRR expected for non-aberrant probes. The dataset contains 5 different regions $R_1, \ldots, R_5$ delimited by vertical green lines.

2˙2.1. The Model. Let m and n be the number of observed samples and the number of probes observed for a chromosome respectively and let $D \in \mathbb{R}^{mn}$ be the data matrix where the element $D_{ij} \in D$ contains the LRR observed on the j-th probes of the sample i, then we define a segmentation S of D as a set of ordered positions (breakpoints) partitioning the columns of D into M connected regions $\mathbf{R} = \{R_1, \cdots, R_M\}$. Here we use the 1-dimensional version of the piecewise constant Mumford and Shah functional, in this case the length of the boundaries between regions has no influence on the segmentation, and the second term of (2) reduces to the number of regions, denoted here as M.

$$E(u, \Gamma) = \sum_i \int_{\Omega_i} (\mathbf{u}_o - \mathbf{u}_i)^2 \mathrm{d}x \; \mathrm{d}y + \lambda M. \tag{3}$$

Note that we are considering that $\mathbf{u}_o$ is a vector-valued function characterized by several channels, where each channel represents a different sample. Optimal segmentation must be chosen among the 2^n possible solutions. In aCGH data we have a resolution that provides observations for more than 1 million of probes, so brute force algorithms cannot be used and heuristic strategies must be applied. Here we use a greedy procedure for minimization of (3).

2˙2.2. Minimization Process. In order to compute the minimization of (3) we use a region growing process where small adjacent regions are progressively merged to create larger ones. In this way use a pyramidal algorithm that moves from fine to coarse segmentations. Given two adjacent regions R_i and R_j the merging criterion used in the minimization process is:

$$E(u, \Gamma \setminus R_i \cup R_j) - E(u, \Gamma)) = \frac{|R_i| \, |R_j|}{|R_i| + |R_j|} \|\mathbf{u}_i - \mathbf{u}_j\|^2 - \lambda, \tag{4}$$

where $R_i, R_j \in \mathbf{R}$ with $i \neq j$, $|R_i|$ is the length of the i-th region and $\|\cdot\|$ represents the weighted L_2 norm. Following a greedy procedure, the minimization process starts with a segmentation having n regions each for each probe and in each step the pair of regions that yields the best score are merged into a single region. In particular, given a fixed value of λ, the algorithm iteratively looks for the pair of adjacent regions for which the criterion in (4) is as negative as possible. If no pair of regions having negative value of (4) exists, then the value of λ is increased. The resulting method is considered a multiscale algorithm [20] where the value of λ represents the *scale*: as λ grows the segmentation

gets coarser. The algorithm stops when the maximum value of the scale parameter is reached. The sequence of values of this parameter is called λ-schedule by analogy with temperature schedule of simulated annealing.

2˙2.3. λ-schedule. λ-schedule is the sequence of λ used in the minimization process. Careful selection of λ-schedule is fundamental to produce a "good" segmentation. From eq. (4) we can deduce that the cost required for merging of two adjacent regions R_{i-1} and R_i is

$$\hat{\lambda}_i = \frac{|R_{i-1}|\,|R_i|}{|R_{i-1}| + |R_i|}\|\mathbf{u}_{i-1} - \mathbf{u}_i\|^2. \tag{5}$$

Given a fixed value $\overline{\lambda}$ representing the current value for the scale parameter in the minimization process, two adjacent regions are merged if the respective cost $\hat{\lambda}_i < \overline{\lambda}$. If no adjacent regions satisfy this inequality, then no merging is allowed and, in order to continue in the minimization process, a new value for $\overline{\lambda}$ must be chosen. Here, as in Vega, the next scale parameter value of the λ-schedule is chosen as the smallest $\hat{\lambda}_i$ (plus a positive constant close to zero). By using this update rule new region merges are allowed, so the region growing process will be composed by fine merging operations. In this way the algorithm follows a greedy strategy where the selection of the next $\hat{\lambda}$ within the λ-schedule represents the greedy choice.

3. – Results

We tested VegaMC on a Gastrointestinal Stromal Tumor (GIST) aCGH dataset. GISTs are the most common mesenchymal tumors of the gastrointestinal tract. We used the data published by [4] where 25 fresh tissue specimens of GISTs were collected and hybridized by Affymetrix Genome Wide SNP 6.0 (GEO identifier GSE20710). Raw data were preprocessed by PennCNV tool [21] obtaining the LRR for about 1.6 millions of probes. In order to investigate the performances of VegaMC we used some well-known cytogenetic aberrations characterizing GIST. In particular, deletions of *KIT* and *PDGFRA* genes (located in cytoband 4q12) were identified as important for tumor phenotype in GIST, and their presence correlates with benefit from target therapy [22]. In addition, loss of cytobands 14q11.2, 22q12.2 and 22q13.31 appears to play an important role in early stage of tumor formation and in late tumor progression [23, 24] and gains of 7p11.2 and 12q15 were confirmed by in situ hybridization [25]. Cytogenetic losses of 1p36.23, 9p21.3, 14q23.1, 13q14.3 and gain of 5q35.3 have also been related with GISTs [26, 4].

The list of the considered cytogenetic aberrations is reported in table II, where in order to distinguish between loss and gain we use the following rule: Let $\boldsymbol{\mu}$ be the vector of the mean for a region, then the correspondent CNA is considered a loss if $\|\boldsymbol{\mu}\| < -0.2$ and a gain if $\|\boldsymbol{\mu}\| > 0.2$. These thresholds reflect the suggestions reported in Vega [8]. The first observation is on the number of identified aberrant cytobands: on a total of 147 aberrant cytobands 111 were identified as losses and 37 as gains. This confirmed the results of the investigation of Chen *et al.* [27]: in GIST losses were more common than gains. From table II we can see that VegaMC was in agreement to the biological knowledge on 75% of the considered cytobands. In addition, we can notice that the weighted mean seems to be a good index to summarize the possible state (loss, normal and gain). Indeed, in each corrected evaluation the biological knowledge on the kind of the aberration is confirmed. VegaMC did not detect aberrations for 1p36.23 ($\|\boldsymbol{\mu}\| = -0.03$), 12q15 ($\|\boldsymbol{\mu}\| = 0.01$) and 22q12.2 ($\|\boldsymbol{\mu}\| = -0.11$) cytobands.

TABLE II. – *List of well-known aberrant cytobands in GIST. For each cytoband we report the expected alteration (loss, gain) and the overlapped target genes with the respective descriptions. For each cytoband we also report the results of VegaMC in terms of weighted mean and assigned alteration.*

(Cytoband)	Expected alteration	$\|\boldsymbol{\mu}\|$	Computed alteration	(Target genes)	Ref.
1p36.23	loss	−0.03	normal	*ENO1*	[23]
4q12	loss	−0.22	loss	*KIT*	[22]
4q12	loss	−0.22	loss	*PDGFRA*	[22]
5q35.3	gain	0.21	gain	*GNB2L1*	[4]
7p11.2	gain	0.31	gain	-	[25]
9p21.3	loss	−0.24	loss	*CDKN2A*, *CDKN2B*	[26]
12q15	gain	0.01	normal	-	[25]
13q14.3	loss	−0.26	loss	*INTS6*	[4]
14q23.1	loss	−0.26	loss	*PPM1A*	[4]
14q32.33	loss	−0.34	loss	*PACS2*	[23]
22q12.2	loss	−0.11	normal	NF2	[24]
22q13.31	loss	−0.30	loss	-	[24]

4. – Conclusions

Biological studies highlighted the importance of CNAs in cancer development and progression. Here we described a new algorithm that performs a joint analysis of copy number data (VegaMC) so that a map of the common CNAs can be obtained. VegaMC is based on a popular image segmentation model and it uses a functional that takes into account accuracy and penalizes complex solutions. Minimization of the functional is performed by a greedy multiscale process where a data-driven strategy is used to define the stopping condition in terms of scale parameter. In order to perform a qualitative assessment we used a recently published aCGH dataset. Results show the ability of VegaMC in detection of well-known cytogenetic mutations. Results also suggest that accuracy can be improved by a suitable choice of the threshold used to distinguish between normal and aberrant chromosomal regions. Finally we want to debate about the algorithmic complexity of VegaMC. Finally we want to debate about the algorithmic complexity of VegaMC. In order to improve the performance of VegaMC, in the minimization process we used an heap structure so that the overall complexity is $O(n(m + \log n))$. This implementation strategy allowed to VegaMC to run in about $4'30''$ on GIST dataset which represents a very interesting execution time.

REFERENCES

[1] SHLIEN A. and MALKIN D., *Genome Med.*, **1** (2009) I:62.
[2] FEUK L., CARSON A. R. and SCHERER S. W., *Nat. Rev. Genet.*, **7** (2006) 85.
[3] LOGDON J. A., LAMONT J. M., SCOTT D. K., DYER S., PREBBLE E., BOWN N., GRUNDY R. G., ELLISON D. W. and CLIFFORD S. C., *Genes Chromosomes Cancer*, **45** (2006) 47.

[4] Astolfi A., Nannini M., Pantaleo M. A., Di Battista M., Heinrich M. C., Santini D., Catena F., Corless C. L., Maleddu A., Saponara M., Lolli C., Di Scioscio V., Formica S. and Biasco G., *Lab. Investi.*, **90** (2010) 1285.
[5] Beroukhim R., Mermel G. H., Porter D. *et al.*, *Nature*, **463** (2010) 899.
[6] Venkatachalam R., Verwiel E. T., Kamping E. J., Hoenselaar E., Görgens H., Schackert H. K., van Krieken J. H., Ligtenberg M. J., Hoogerbrugge N., van Kessel A. G. and Kuiper R. P., *Int. J. Cancer*, (2010) .
[7] Olshen A. B., Venkatraman E. S., Lucito R. and Wigler M., *Biostatistics*, **5** (2004) 557.
[8] Morganella S., Cerulo L., Viglietto G. and Ceccarelli M., *Bioinformatics*, **26** (2010) 3020.
[9] Aguirre A. J., Brennan C., Bailey G., Sindhi R., Feng B., Leo C., Zhang Y., Zhang J., Gans J. D., Bardeesy N., Cauwels C., Cordon-Cardo C., Redston M. S., DePinho R. A. and Chin L., *Proc. Natl. Acad. Sci. U.S.A*, **101** (2004) 9067.
[10] Beroukhim R., Getz G., Nghiemphu L., Barretina J., Hsueh T., Linhart D., Vivanco I., Lee J. C., Huang J. H., Alexander S., Du J., Kau T., Thomas R. K., Shah K., Soto H., Perner S., Prensner J., Debiasi R. M., Demichelis F., Hatton C., Rubin M. A., Garraway L. A., Nelson S. F., Liau L., Mischel P. S., Cloughesy T. F., Meyerson M., Golub T. A., Lander E. S., Mellinghoff I. K. and Sellers W. R., *Proc. Natl. Acad. Sci. U.S.A*, **104** (2007) 20007.
[11] Sanchez-Garcia F., Akavia U. D., Mozes E. and Pe'er D., *BMC Bioinformatics*, **11** (2010) 189.
[12] Morganella S., Pagnotta S. M. and Ceccarelli M., *Bioinformatics*, (2011) .
[13] Rueda O. M. and Diaz-Uriarte R., *BMC Bioinformatics*, **10** (2009) 308.
[14] Taylor B. S., Barretina J., Socci N. D., Decarolis P., Ladanyi M., Meyerson M., Singer S. and Sander C., *PLoS One*, **3** (2008) e3179.
[15] Yang W. L., Chipman H. A., Bull S. B., Briollais L. and Wang K., *Bioinformatics*, **25** (2009) 1669.
[16] Pique-Regi R., Ortega A. and Asgharzadeh S., *Bioinformatics*, **25** (2009) 1223.
[17] Klijn C., Holstege H., de Ridder J., Liu X., Reinders M., Jonkers J. and Wessels L., *Nucl. Acids Res.*, **36** (2008) e13.
[18] Ivakhno S. and Tavaré S., *Bioinformatics*, **26** (2010) 1395.
[19] Mumford D. and Shah J., *Commun. Pure Appl. Math.*, **41** (1989) 577.
[20] Koepfler G., Lopez C. and Morel M., *IAM J. Numer. Anal.*, **31** (1994) 282.
[21] Wang K., Li M., Hadley D., Liu R., Glessner J., Grant S. F., Hakonarson H. and Bucan M., *Genome Res.*, **17** (2007) 1665.
[22] Yang J., Du X., Lazar A. J., Pollock R., Hunt K., Chen K., Hao X., Trent J. and Zhang W., *Cancer*, **113** (2008) 1532.
[23] Ässämäki R., Sarlomo-Rikala M., Lopez-Guerrero J. A., Lasota J., Andersson L. C., Llombart-Bosch A., Miettinen M. and Knuutila S., *Genes Chromosomes Cancer*, **46** (2007) 564.
[24] Lasota J., Wozniak A., Kopczynski J., Dansonka-Mieszkowska A., Wasag B., Mitsuhashi T., Sarlomo-Rikala M., Lee J. R., Schneider-Stock R., Stachura J., Limon J. and Miettinen M., *Lab. Invest.*, **85** (2005) 237.
[25] Tornillo L., Duchini G., Carafa V., Lugli A., Dirnhofer S., Di Vizio D., Boscaino A., Russo R., Tapia C., Schneider-Stock R., Sauter G., Insabato L. and Terracciano L. M., *Lab. Invest.*, **85** (2005) 921.
[26] Perrone F., Tamborini E., Dagrada G. P., Colombo F., Bonadiman L., Albertini V., Lagonigro M. S., Gabanti E., Caramuta S., Greco A., Torre G. D., Gronchi A., Pierotti M. A. and Pilotti S., *Cancer*, **4** (2005) 159.
[27] Chen Y., Tzeng C. C., Liou C. P., Chang M. Y., Li C. F. and Lin C. N., *J. Biomed. Sci.*, **11** (2004) 65.