Oral cancer develops and progresses by accumulation of genetic alterations. The interrelationship between these alterations and their sequence of occurrence in oral cancers has not been thoroughly understood. In the present study, we applied oncogenetic tree models to comparative genomic hybridization (CGH) data of 97 primary oral cancers to identify pathways of progression. CGH revealed the most frequent gains on chromosomes 8q (72.4%) and 9q (41.2%) and frequent losses on 3p (49.5%) and 8p (47.5%). Both mixture and distance-based tree models suggested multiple progression pathways and identified +8q as an early event. The mixture model suggested two independent pathways namely a major pathway with −8p and a less frequent pathway with +9q. The distance-based tree identified three progression pathways, one characterized by −8p, another by −3p and the third by alteration +11q and +7p. Differences were observed in cytogenetic pathways of node-positive and node-negative oral cancers. Node-negative cancers were characterized by more non-random aberrations (n = 11) and progressed via −8p or −3p. On the other hand, node-negative cancers involved fewer non-random alterations (n = 6) and progressed along −3p. In summary, the tree models for oral cancers provided novel information about the interactions between genetic alterations and predicted their probable order of occurrence.

Key words: oral cancer; CGH; comparative genomic hybridization; oncogenetic tree; progression pathways; genetic progression score

For each cancer type, the tree models have identified multiple progression pathways and revealed different subtypes characterized by combinations of alterations.

To date, the pathogenetic pathways followed by oral cancer have not been thoroughly investigated. Elucidating the divergent routes in oral carcinomas could provide information about molecular subtypes, which might support treatment decisions. With this long-term goal in mind, in the current study, we constructed oncogenetic tree models based on CGH data of 97 primary oral cancers. Both mixture models and distance-based trees models were used to analyze genetic alterations in OSCC. Distance-based tree models were also constructed separately for node-positive and node-negative OSCC.

Material and methods

Study population

Tumor tissues were collected from 97 oral cancer patients who underwent surgical resection at the Tata Memorial Hospital (TMH), Mumbai. Tissue collection and the entire study protocol were approved by the Institutional Review Board at TMH. Informed consent was obtained from the patients. All patients underwent neither chemotherapy nor radiation therapy before surgery. After microdissection of tissues, the pathologist confirmed that each tissue sample had ≥60% tumor cell content. All the tumor samples were graded and staged according to the WHO and TNM and AJCC 2002 classification of tumors, respectively. Clinicopathological data of the cases are summarized in Table I. The study group consisted of 75 male and 22 female with median age of 52 years, ranging from 23 to 77 years. There were more node-positive tumors (n = 54) than node-negative tumors (n = 43).

CGH analysis

CGH was performed as described previously using direct labeling method with fluorochrome labeled dUTPs. Using standard nick translation method, tumor DNA was labeled with Fluorescein-12-dUTP and normal DNA was labeled with Texas Red-5-dUTP (NEN, Boston, MA). An equal quantity (2 μg each) of tumor and normal DNA were mixed with 10 μg of unlabeled human Cot-1 DNA (Invitrogen) and dissolved in 10 μl of hybridization buffer. The denatured probes in the mixture were hybridized onto the normal metaphase spreads (Vysis, Touhy, Des Plaines, IL) at 37°C for 48 hr. After post-hybridization washes, the slides were

Grant sponsor: Indian Council of Medical Research (ICMR); Grant number: 5/132/TF/2001-NCD-III. Grant sponsor: Intramural Research program of the National Institutes of Health.

*Correspondence to: Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Cancer Research Institute (CRI), Tata Memorial Centre (TMC), Kharghar, Navi Mumbai-410210, India. Fax: +91-22-27405085.
E-mail: mmahimkar@actrec.gov.in or mahimkarm@yahoo.in
Received 14 September 2008; Accepted after revision 12 December 2008
DOI 10.1002/ijc.24267
Published online 13 January 2009 in Wiley InterScience (www.interscience.wiley.com).

Int. J. Cancer: 124, 2864–2871 (2009)
© 2009 UICC

Construction of oncogenetic tree models reveals multiple pathways of oral cancer progression

Swapnali Pathare1, Alejandro A. Schäffer2, Niko Beerenwinkel3 and Manoj Mahimkar1*1

1Advanced Centre for Treatment, Research and Education in Cancer (ACTREC), Cancer Research Institute (CRI), Tata Memorial Centre (TMC), Kharghar, Navi Mumbai, India
2Computational Biology Branch, National Center for Biotechnology Information, NIH, DHHS, Bethesda, MD
3Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

72
counterstained with DAPI (Vectorshied, Burlingame, CA) for chromosome identification before visualization by the fluorescence microscopy (Zeiss Axioscope, Zena, Germany). The images were analyzed by digital image analysis system (Metasystems, Germany). The average ratio of green to red fluorescence intensity was calculated for each chromosome. The thresholds were set at 1.25 and 0.75 to determine copy number alterations (CNAs), i.e., gains and losses, respectively.

The thresholds 1.25 and 0.75 are commonly, though not universally, used in CGH. The thresholds were proposed in the original article on CGH\(^{21}\) and further evaluated in early studies.\(^{22,23}\) The study of Jeuken et al.\(^{24}\) showed that using thresholds of 1.2 and 0.8 did not make much difference in practice. Among the 6 previous studies on CGH of oral cancer that we found,\(^ {4–9}\) some used 1.25/0.75 thresholds; others used 1.2/0.8. As far as we could determine, none of these 6 studies analyzed their data using more than one pair of thresholds.

### Fluorescence in situ hybridization

Interphase FISH was performed on archival oral cancer samples with known CGH results to evaluate the CGH-predicted gains of 11q13 and 8q24.3. Dual colored FISH was performed on 4 μm sections of archival OSCC samples (n = 28) and the corresponding normal oral mucosal tissues (n = 3). Centromere-specific BAC clones were Cy-3 (red) labeled and region-specific BAC clones were FITC (green) labeled using standard nick translation method. All BAC clones were obtained from CHORI (BACPAC Resource Center, Children’s Hospital and Research Centre, Oakland, USA). The specific BAC clones we used were RP11-642A1 for region 8q24.3, RP11-73M19 for chromosome 8 centromere, RP11-149G19 for 11q13 region and RP11-135H8 for chromosome 11 centromere.

Sections on slides were first deparaffinized in xylene and then treated using commercially available Vysis pre-treatment kit (Vysis). Briefly, this involved treating the sections for 30 min in 1 M sodium thiocyanate, digesting for 20 min with protease at 0.5 mg/ml in 0.01 N HCl and fixing in 10% neutral buffered formalin. For FISH experiments, labeled probes were added to a slide in a hybridization solution containing 50% deionized formamide, 10% dextran sulphate, 2X SSC, 2 mg salmon sperm DNA and 10 mg Cot-1 DNA. The slides and probe DNA were denatured at 75°C for 10 min and hybridized overnight in a humidified chamber at 37°C. Subsequently, the slides were subjected to post-hybridization washes 50% deionized formamide/2X SSC at 42°C for 10 min and three times in 2X SSC at 42°C for 5 min. Interphase nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI).

For evaluation of the experiments, hybridization signals from 100 non-overlapping interphase cell nuclei of each tumor sample were counted using a fluorescence microscope. A copy number gain was scored, if the average number of signals per nucleus was greater than or equal to three (≥3).

### Construction of oncogenetic trees

Two types of oncogenetic tree models were used to describe the occurrence of genetic alterations during the progression of oral cancer. Distance-based trees were constructed using the software oncotrees\(^ {24,25}\) (http://www.ncbi.nlm.nih.gov/CRCResearch/Schaffer/cgh.html) and mixture models of trees were generated by the Mtremix software package (http://mtremix.bioinf.mpi-sb.mpg.de).\(^ {25}\)

Chromosomal aberrations from CGH data were used as input for the tree modeling procedures. Abnormalities on 1p, 16p, 22q and Y chromosome arms were excluded from analysis, because these chromosome regions are guanine/cytosine rich regions, which are known to yield false positives. For construction of tree models, the CGH profile was recorded as presence/absence of a gain and presence/absence of a loss on a chromosome arm. We first tried using more precise single digit bands (data not shown), but this representation led to confounding of spatial relationships between bands on the same chromosome arm and the desired temporal relationships between copy number aberrations.

Trees were constructed from CNA events selected as non-random by the method of Brodeur et al.\(^ {26}\) which is implemented within oncotrees. The method of Brodeur et al. requires a prior distribution for CNA occurrences. For this purpose, we initially assumed that the probabilities of a gain or loss on a chromosome arm are equal and proportional to the chromosome arm size; we used arm sizes from Morton\(^ {27}\). However, it is known that gains are more common than losses in OSCC\(^ {28,29}\) and in the present data set, gains were approximately twice as frequent as losses. Therefore, we also used a 2:1 skewed prior distribution in addition to the balanced prior. To reduce the number of events selected as non-random, we used a threshold of 99th percentile in the test statistic of Brodeur et al. The events selected as non-random scored above the 99th percentile using both the balanced and the 2:1 skewed prior distributions. The selection of non-random events was redone for each tumor subset considered.

In the oncogenetic tree model, the root corresponds to the normal state of the cell. Vertices of the tree represent genetic alterations (events) and edges between vertices represent statistical dependencies between events. Each vertex is associated with the probability that the corresponding event will occur if the preceding event in the tree has already occurred. Thus, aberrations that are placed close to the root of the tree are estimated to occur early in tumor development, whereas those at longer distances are estimated to occur late in tumor progression. Correlated events tend to cluster in sub-trees. In the tree models, the genetic events are assumed to be irreversible.

### Mixtures of oncogenetic trees

The oncogenetic trees mixture model\(^ {12,18,25}\) consists of several weighted components, each of which is an oncogenetic tree as described above. Usage of several tree components allows more flexible modeling of oncogenetic pathways than using a single tree. We fixed one tree component with a star topology and uniform transition probabilities. This tree models events as independent and uniform and thus serves as a null model. The fraction of samples assigned to this tree does not show apparent branching-like correlations. By contrast, the structure and parameters of the other tree components are learned from the data and they reveal pronounced dependencies among the cytogenetic alterations. The star component captures the possibility that gains and losses occur at random with no dependencies. The mixture model allows the addition of one or more non-star trees to model the dependencies

### Table 1 - CLINICO-PATHOLOGICAL CHARACTERISTICS OF 97 ORAL CANCER PATIENTS

| Characteristics | No. of patients (%) |
|-----------------|---------------------|
| Gender          |                     |
| Male            | 75 (77)             |
| Female          | 22 (23)             |
| Median age      | 52 yr (range, 23–77)|
| <52 yr          | 46 (47)             |
| >52 yr          | 51 (53)             |
| Tumor sites     |                     |
| BM and allied sites\(^1\) | 69 (71) |
| Tongue          | 28 (29)             |
| Differentiation |                     |
| Well            | 14 (14)             |
| Moderate        | 53 (55)             |
| Poor            | 30 (31)             |
| Lymph node involvement |             |
| Negative (pN0)  | 43 (44)             |
| Positive (pN+)  | 54 (56)             |
| Stage (pTNM)    |                     |
| I and II        | 14 (14)             |
| III and IV      | 83 (86)             |

\(^1\)BM, buccal mucosa; allied sites, alveolus, gingiva, retromolar trigone.
between copy number aberrations. Using more non-star trees may fit the data better, but is susceptible to over fitting. We used a modified Bayesian Information Criterion (BIC) to trade off the increasing complexity of the model with the improved fit to the data as the number of tree components increases.28 Using this model selection technique, one additional non-star tree gave the best BIC score.

Distance-based trees

The distance-based trees were constructed as described in previous studies.16,29 For the distance-based trees, we use the Fitch30 and Neighbor,31 programs of PHYLIP32 to fit the distances to a tree. In all cases, the Neighbor program gave a better fit, which means the following. Let $I_{ij}$ be the input distance between CNAs $i$ and $j$ and let $T_{ij}$ be the distance implied by the output tree. We have two choices for the matrix $T$, one from the Fitch program and other from the Neighbor program. We define a matrix of differences $D_{ij} = |I_{ij} - T_{ij}|$ one choice of $T$ “fits” better than another, if it gives a matrix $D$ that has smaller entries. The total size of $D$ can be measured by standard matrix norms. When we do so, for these trees, the tree distances $T$ from Neighbor always give a better fit than those from Fitch.

To compute bootstrap confidence levels for each split in the distance-based trees we used a 3-step procedure. First, 100 bootstrap samples were generated and their associated distance matrices were computed using the bootstrapping module of oncetrees. Second, we used the Neighbor program to fit each bootstrap sample to a tree. Third, we used the Consense program from PHYLIP to compute the number of times each split in the tree occurred perfectly. This method counts positively only those trees in which a split into sub-trees occurred exactly as in the original distance-based tree; therefore, one expects confidence levels to be fairly low when the split involves more than 2 or 3 events.

Tree visualization

Mixture models are drawn with Graphviz.33 and distance-based trees are drawn with TreeView.34 The relevant distance in the visualization of distance-based trees is the horizontal distance from the root node to the node representing a copy number aberration or the horizontal distance between 2 copy number aberrations. Vertical edges and distance are used to spread out the tree for easier visualization.

Genetic progression scores

The genetic progression score (GPS) of an observed tumor sample is defined as the expected waiting time of the mutational pattern of the tumor in the timed oncogenetic trees mixture model.18 The absolute values of this quantity are only meaningful if information on the true age of each tumor is available. However, even without scaling, the expected waiting time provides a useful measure of genetic progression that is based on the tree model. Unlike simple counting of alterations, the GPS accounts for dependencies between events. The GPS distributions of node-positive and node-negative patients were compared using the Wilcoxon rank-sum test.

Time of occurrence

A simpler method to infer the relative early and late occurrence of the events is called time of occurrence (TO) analysis.10,35 One computes, for each event, how many other events occur in all tumors that have the event. The general concept of TO analysis is that an event A occurs before an event B if the number of events co-occurring with A is smaller than with B. To compare the number of co-occurring events for A and B, one could use test statistics the average, the median, or the mode. Desper et al.10 recommended using the average. Höglund et al.33 who coined the term “TO analysis”, recommended using the mode. Unfortunately, in our data set, many CNAs have a mode of “9” in the distribution of number of co-occurring aberrations, so the mode is not a very useful test statistic. Therefore, we followed the recommendation of Desper et al.10 and used the average number of co-occurring events to propose an order of events.

Results

Copy number alterations detected by CGH in 97 primary oral cancers

Copy number gains were more common than losses. The average number of total CNAs was 9.90 (range, 1–30). The most common gains were observed on chromosome regions 8q (74.2%), 9q (41.2%), 11q13 (41.2%), 7p (37.1%), 3q (35.1%), 20q (33%), 20p (29%) and 5p (24%). The most common losses were observed on chromosome regions 3p (49.5%), 8p (47.5%), 18q (34%) and 11q14-qter (20.6%). The bandless CGH data are available at: ftp://ftp.ncbi.nlm.nih.gov/pub/oral_cancer.

Confirmation of some CGH results by FISH

To confirm some of the CNAs detected by CGH, we performed FISH on 28 samples to evaluate the regions 11q13 and 8q24 that are frequently gained (Fig. 1). For samples where we could obtain FISH signals, the concordance rates were good (81% for 11q13 and 88% for 8q24; Table II). For 11q13, the 5 discrepancies were due to tissues that were non-informative in 2 cases and FISH detected a gain that was not found by CGH in 3 cases. For 8q24, the 3 discrepancies comprise 2 non-informative tissues and one case of FISH detected a gain that was not found by CGH.

Construction of oncogenetic trees for oral cancer progression

The method of Brodeur et al. (see Material and Methods) selected 12 CNAs that occurred more frequently than would be expected at random: +5p, +7p, +8q, +9q, +11q, +17p, +18p, +20p, +20q, −3p, −8p and −18q. Here, a plus (+) symbol indicates the gain of a chromosomal region and a minus (−) represents a loss. Oncogenetic and distance-based trees were constructed using these 12 events. Both models represent the apparent multi-step and multi-pathway process of oral carcinogenesis (Figs. 2 and 3).

Oncogenetic trees mixture model

We estimated oncogenetic trees mixture models consisting of a star component and a non-star tree component to obtain a concise description of the genetic development of oral cancers (Fig. 2). A third of the tumors can be explained by the non-star tree component. In this branching tree, the root vertex corresponds to the normal oral keratinocyte, whereas the other vertices represent the CNAs of interest. Because the event +8q is the only direct successor of the root, it is predicted to be the initial event. Once this event occurs, the occurrence of subsequent events becomes much more likely. After +8q, the branching tree displays two independent pathways, one consisting only of event +9q, the other comprising the 10 remaining events and starting with −8p followed by −3p. The latter pathway was predicted to further branch into two pathway beginning with −18q and +7p, respectively. After the initial +8q event, the large 10-event sub-branching is more likely to develop than the +9q pathway, the likelihood ratio being 0.79:0.51 = 1.55.

Distance-based trees

In the distance-based tree model, the time to occurrence is proportional to the root–leaf horizontal distance (Fig. 3). According to this model, +8q was an early event. After the occurrence of +8q, the distance-based tree classified other events into 2 or 3 clusters. One cluster is marked by −8p and −3p (comprising −8p, −18q and +18p; and −3p, +9q and +17q); it might be split into two sub-clusters of three events each, but the bootstrap confidence for the split into two sub-clusters is low (14%). The other cluster is marked by events +11q and +7p (comprising +11q, +7p, +20p, +20q and +5p). These two clusters suggest OSCC genetic subtypes.

Though similarities were observed among the tree models, there were inconsistencies in terms of whether (i) −3p is part of the −8p pathway and (ii) +9q depends on −3p or not.
Construction of oncogenetic trees for node-negative and node-positive oral cancers

In a univariate analysis comparing node-positive and node-negative tumors, four events showed significant association with node-positive status by a one-sided Fisher’s exact test: 2p (p < 0.008), 17p (p < 0.01), 218q (p < 0.04) and 19q (p < 0.04). These associations do not remain significant after correcting for multiple testing using the false discovery rate (FDR) method.36 Using the FDR method, 2p (adjusted p < 0.06) and 7p (adjusted p < 0.06) approached statistical significance. Thus, associations of single genetic imbalances with nodal status are not significant after correcting for multiple testing. However, this finding does not preclude the possibility that associations of 2 or more imbalances with nodal status might be statistically significant. To test this hypothesis, we focused on multivariate analyses, which could shed more light on the genomic differences between node-positive and node-negative OSCC.

Separate distance-based tree models were constructed to understand the progression pathways in node-positive versus node-negative oral cancers. More events (n = 11) were selected as non-random in node-positive OSCC than in node-negative OSCC (n = 6) (Fig. 4), consistent with the hypothesis that node-positive tumors are more advanced. For both sets of tumors, the non-random events were chosen systematically by the established method of Brodeur et al. (1982), although we raised the cutoff to the 99th percentile instead of the 95th. Being more stringent is not helpful. If we are less stringent (e.g., 95th percentile), then the number of events selected as non-random grows and the tree models become unwieldy.

According to Figure 4, the node-positive cancers may be classified into two main groups: one group includes +9q, −8p, +18p and −18q, whereas the other group includes, +7p, +11q, +20p, +20q and +5p. Events +8q and −3p are early events that do not fit clearly into either of these 2 groups. The splitting of the 2 large groups has only moderate bootstrap support (15%), whereas the separation of +8q and −3p is more pronounced (63% and 25%, respectively). The event +9q, which is suggested as an important
indicator of progression by the mixture tree model on all tumors, is selected as non-random for the node-positive tumors, but not for the node-negative tumors. The numbers of tumors in each subset were considered too small for mixture tree analysis.

Further, the distance-based tree models predicted progression of node-negative OSCC by two main classes (Fig. 4, bootstrap confidence 69%). One class included tumors that did not progress much beyond the aberrations on chromosome 8. The other class had 2 subgroups for which the confidence in the split is 64%: subgroup 1, which included tumors that progressed with 8p, 3p and 11q. The events in both these clusters were not placed near to each other. The alterations 7p, 18q, 20p and 5p were not selected for the node-negative OSCC. Overall, the distance-based tree model indicates that this subtype was karyotypically less complex as compared with the node-positive OSCC.

**Genetic progression score**

For each observed tumor, the GPS measures the level of genetic progression in the oncogenetic tree model. Node-positive cancers had progressed significantly further along the tree model than node-negative cancers (median GPS of 0.87 vs. 0.59, p = 0.0016; Fig. 5).

**Time of occurrence**

TO statistics for the non-random events are shown in Table III. The average number of imbalances in tumors harboring −8p was less than the average for −3p, which indicates that −8p would...
occur before \(-3p\). The values for all alterations except \(+18p\) were consistent with the predictions of the tree analysis. The exception might be due to the lower frequency of \(+18p\) in our study (13.4%).

**Discussion**

The non-random CGH pattern of oral cancers detected in our study was generally consistent with previous studies,\(^4\)–\(^9\) although some of the losses reported as common vary from study to study; this variation may be due to differences between the populations sampled.\(^9\) Among the observed alterations, the gain of 8q and loss of 3p have been reported to be the early events in oral carcinogenesis.\(^38\),\(^39\) Besides these alterations, the sequence of other aberrations in oral tumorigenesis is not known. Elucidating progression pathways from complex CGH data requires statistical
modeling techniques such as oncogenetic trees. The current study is first to apply tree models to CGH data obtained from a reasonably large number (n = 97) of primary oral cancers. Among the previous CGH studies of oral cancer, the largest number of tumors in one study was 35.16

The branching mixture models and distance-based tree models identified divergent pathways of progression in oral cancers. There were similarities and differences in the sequence of alterations revealed by them. The common findings led to the following inferences: (i) +8q is located near the root, hence it is an early event in cellular transformation of oral cancers; (ii) at least three subtrese emerge subsequent to the occurrence of +8q, which indicate that there are divergent pathways of progression in oral carcinomas; (iii) alterations +p5, +17p and +18p are late events due long distance from the root; (iv) a close relationship exist between +p7 and +11q; and +20p and +20q alterations, hence these are present in the same subcluster. Some inconsistencies were observed mainly in the relative ordering of −p3 and +9q alterations, namely whether −p3 is part of the −8p pathway or not and with respect to the placement of +9q, whether it is part of the −3p pathway or independent of the other events. Part of these inconsistencies may result from the fact that only about a third of the data can be mapped to the oncogenetic tree component displayed in Figure 2.

Comparing the current tree models of oral cancers with the tree analysis of the head and neck cancers, we found that some of the alterations (−p3, +8q, +5p, +17p and +18p) were selected as non-random events for both the present oral cancer data set and an early collection of head and neck cancer.16 Also, the prediction of occurrence of +8q and +18p suggested for the subset of head and neck cancers, coincided with the progression pathways identified by our study. Thus, +8q appears to be an early event in the development of oral cancers. Activation of the oncogene c-MYC on chromosome region 8q24 has frequently been implicated in oral carcinogenesis, but some studies have suggested other genes on 8q.17

After the occurrence of +8q, the losses of 8p and 3p were predicted as subsequent early events during progression of OSCC. This novel finding seems to be in contrast to previous studies, which might help elucidating the pathogenesis of oral cancers.

Acknowledgements

The authors thank the Council of Scientific and Industrial Research (CSIR) for providing fellowship to Ms. Swapnali Pathare during her tenure as a graduate (PhD) student. We thank Dr. R. Mistry, Dr. A.K. D’Cruz and Dr. K.A. Pathak who permitted collection of oral cancer samples from TMH; Dr. A.M. Borges for microdissection of tumor samples and Ms. Sadhana Kamman for her valuable suggestions during data analysis. This work has been approved by the Institutional Ethics committee (Project No. 69 of 2001).

References

1. Baudis M. Genomic imbalances in 5918 malignant epithelial tumors: an explorative meta-analysis of chromosomal CGH data. BMC Cancer 2007;7:226.

2. Hoglund M, Gisselsson D, Mandahl N, Johanson B, Mertens F, Mittelman F, Säll T. Multivariate analyses of genomic imbalances in solid tumors reveal distinct and converging pathways of karyotypic evolution. Genes Chromosomes Cancer 2001;31:156–71.

3. Shingaki S, Takada M, Sasai K, Bibi R, Kobayashi T, Nomura T, Saito C. Impact of lymph node metastasis on the pattern of failure and survival in oral carcinomas. Am J Surg 2003;185:278–84.

4. Hermens MAJA, Joene H, Arwert F, Braakhuis BJM, Baak JP, Westerveld A, Slater R. Assessment of chromosomal gains and losses in oral squamous cell carcinoma by comparative genomic hybridisation. Oral Oncol 1997;33:414–18.

5. Wolff E, Giordi S, Liehr T, Vorderwulbecke U, Ries J, Steinsinger H, Gebhart E. Oral squamous cell carcinomas are characterized by a rather uniform pattern of genomic imbalances detected by comparative genomic hybridisation. Oral Oncol 1999:34:186–90.

6. Weber RG, Scheer M, Born IA, Joos S, Cobbers JMIL, Hofele C, Reifenberger G, Zöller JE, Lichter P. Recurrent chromosomal imbalances detected in biopsy material from oral premalignant and malignant lesions by combined tissue microdissection, universal DNA amplification and comparative genomic hybridization. Ann J Pathol 1998;155:295–303.

7. Okafuji M, Ita M, Oga A, Hayatsu Y, Matsuo A, Shinzato Y, Shinzaki F, Sasaki K. The relationship of genetic aberrations detected by comparative genomic hybridization to DNA ploidy and tumor size in human oral squamous cell carcinomas. J Oral Pathol Med 2000;29:226–31.

8. Lin S-C, Chen Y-J, Kao S-Y, Hsu M-T, Lin C-H, Yang S-C. Chromosomal changes in becel-associated oral squamous cell carcinomas and their relationship to clinical parameters. Oral Oncol 2002;38:266–73.

9. Noutomi Y, Oga A, Uchida K, Okafuji M, Ita M, Kawauchi S, Furuya T, Ueyama Y, Sasaki K. Comparative genomic hybridization reveals genetic progression of oral squamous cell carcinoma from dysplasia via two different tumorigenic pathways. J Pathol 2006;206:67–74.
17. Huang Z, Desper R, Kallioniemi O-P, Moch H, Papadimitriou CH, Schäffer AA. Inferring tree models for oncogenesis from comparative genome hybridization data. J Comput Biol 1999;6:37–51.

18. Huang Q, Yu GP, McCormick SA, Däumer M, Hoffmann D, Kaiser R, Schäffer AA. Distance-based reconstruction of tree models for oncogenesis. J Comput Biol 2000;7:789–803.

19. Beerenwinkel N, Rahnenführer J, Däumer M, Hoffmann D, Kaiser R, Schäffer AA. Learning multiple evolutionary pathways from cross-sectional data. J Comput Biol 2005;12:584–98.

20. Jetter R, Urbschat S, Henn W, Feiden W, Beerenwinkel N, Lengauer T. Estimating cancer survival and analysis of tree models for chromosomal classification of diffuse large B-cell lymphomas. J Mol Med 2007;75:801–14.

21. Lichter P, Cremer T. Quantitative analysis of comparative genomic hybridization. Cytometry 1995;19:27–41.

22. Bockmühl M, Gisselsson D, Sill T, Mittelmann F. Coping with complexity: multivariate analysis of tumor karyotypes. Cytogenet Genome Res 2002;105:103–9.

23. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc B 1995;57:289–300.

24. Paterson IC, Eveson JW, Prime SS. Molecular changes in oral cancer may reflect aetiology and ethnic origin. Eur J Cancer B Oral Oncol 1996;32B:150–3.

25. Garnis C, Coe BP, Ishkanian A, Zhang L, Rosin MP, Lam WL. Novel regions of amplification on 8q distinct from the MYC locus and frequently altered in oral dysplasia and cancer. Genes Chromosomes Cancer 2004;39:93–9.

26. Roz L, Wu CL, Porter S, Scully C, Speight P, Read A, Sloan P, Thaker N. Allelic imbalance on chromosome 3p in oral dysplastic lesions: an early event in oral carcinogenesis. Cancer Res 1996;56:1228–31.

27. Rizvi KA, Priddy R, Le ND, Zhang L. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. Clin Cancer Res 2000;6:357–62.

28. Gehbarto E, Ries J, Willfanger J, Lietz R, Effert T. Genomic gain of the epidermal growth factor receptor harboring band 7p12 is part of a complex pattern of genomic imbalances in oral squamous cell carcinomas. Arch Med Res 2004;35:385–94.

29. Kalyankrishna S, Grandis J. Epidermal growth factor receptor biology in head and neck cancer. J Clin Oncol 2006;24:2666–72.

30. Michalides R, van Veenen N, Hart A, Loftus B, Wijntjens E, Balm A. Overexpression of cyclin D1 correlates with recurrence in a group of forty-seven operable squamous cell carcinomas of the head and neck. Cancer Res 2000;60:975–8.

31. Rodrigo JP, García LA, Ramos S, Lazo PS, Suárez C. EMS1 gene amplification correlates with poor prognosis in squamous cell carcinoma of the head and neck. Clin Cancer Res 2000;6:3177–82.

32. Bockmühl U, Iswad CS, Ferrell RE, Gollin SM. Association of 8p23 deletions with poor survival in head and neck cancer. Otolaryngol Head. Neck Surgery 2001;124:451–55.

33. Pearlstein RP, Benninger MS, Carey TE, Zarbo RJ, Torres FX, Rybicki BA, Dyke DL. Loss of 18q predicts poor survival of patients with squamous cell carcinoma of the head and neck. Genes Chromosomes Cancer 1998;21:333–9.