Multiple Microalterations Detected at High Frequency in Oral Cancer

Corisande Baldwin,

Received for reprints: Corisande Baldwin, British Columbia Cancer Research Centre, 675 West 10th Avenue, Vancouver, British Columbia, Canada. Phone: 604-675-8112; Fax: 604-675-8185; E-mail: cbaldwin@bccrc.ca.

Requests for reprints: Corisande Baldwin, British Columbia Cancer Research Centre, 675 West 10th Avenue, Vancouver, British Columbia, Canada, V5Z 1L3. Phone: 604-675-8112; Fax: 604-675-8185; E-mail: cbaldwin@bccrc.ca.

© 2005 American Association for Cancer Research.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Abstract

The development of array comparative genomic hybridization (array CGH) at tiling-path resolution has enabled the detection of gene-sized segmental DNA copy number gains and losses. Here, we present the first application of whole genome tiling-path array CGH to archival clinical specimens for the detailed analysis of oral squamous cell carcinomas (OSCC). We describe the genomes of 20 OSCCs as well as a selection of matched normal DNA in unprecedented detail. Examination of their whole genome profiles enabled the identification of alterations ranging in size from whole-arm, segmental, to gene size alterations. Tiling-path resolution enabled the detection of many more alterations within each tumor than previously reported, many of which include narrow alterations found to be frequent events among the 20 OSCCs. We report the presence of several novel frequent submegabase alterations, such as the 0.58 Mb gain at 5p15.2 containing triple functional domain (TRIO), detected in 45% of cases. We also report the first coamplification of two gene clusters, by fine-mapping the precise base pair boundaries of the high-level amplification at 11q22.2-22.3 containing both matrix metalloproteinase and baculoviral IAP repeat-containing protein 2 (BIRC) gene clusters. These results show the large improvement in detection sensitivity and resolution compared with genome interval marker arrays and the utility of tiling resolution array CGH for the detection of both submegabase and single copy gains and losses in cancer gene discovery. (Cancer Res 2005; 65(17): 7561-7)

Introduction

Oral cancer is the most common head and neck neoplasm, affecting >400,000 people worldwide each year. The prevention and management of this disease will greatly benefit from the identification of molecular markers and targets (1, 2). Frequent loss of heterozygosity or recurrent segmental DNA copy number changes are indicative of chromosomal regions containing genes critical to tumorigenesis. Array comparative genomic hybridization (array CGH) enables the detection of segmental gains and losses of DNA. We previously reported the application of this technology to identify genetic alterations on chromosome arm 3p in oral carcinoma (3). The recent development of a bacterial artificial chromosome (BAC) tiling-path array that spans the human genome with 32,433 overlapping clones has made the analysis of the entire genome possible at unprecedented resolution (4), revealing both large-scale as well as gene-size alterations, which are not visible using conventional megabase interval genome-wide array CGH analysis such as that described by Snijders et al. (5). Here, we present the first application of whole genome tiling-path array CGH to clinical specimens for the detailed analysis of oral squamous cell carcinoma (OSCC) genomes in order to identify numerous novel frequent submegabase alterations.

Materials and Methods

Tissue samples. Formalin-fixed paraffin-embedded OSCCs were obtained from the British Columbia Oral Biopsy Service and diagnoses confirmed as part of the quality control of array production. This involves hybridization of normal male versus normal female DNA and examination of the X chromosome profiles to control for detection sensitivity of single copy change as previously described (4).

Sample and reference (normal diploid male) genomic DNA (400 ng each) was random prime labeled with 4 nmol of cyanine-3 and cyanine-5 dCTP, respectively. Purified samples were hybridized and washed as previously described (3). Spot images were captured and analyzed using an Arrayworx scanner and SoftWorx Tracker Spot Analysis software (Applied Precision, Issaquah, WA). SeeGH custom software was used to visualize all data as log2 ratio plots (7). This software is available publicly online. Clones with SDs between triplicate spots >0.075 or with signal-to-noise ratios of <10 were stringently filtered from the raw data.

Genetic alterations and their associated breakpoints were identified using both aCGH-Smooth software (8) and visual analysis. The aCGH-Smooth software package identifies copy number gains and losses using a local search algorithm that calculates the probability that the signal ratio for each BAC clone corresponds with the same copy number status as a set of nearby clones using maximum likelihood estimation (8). Visual analysis of the normalized data in SeeGH by multiple karyogram alignment was employed to confirm alterations identified by aCGH-Smooth and to define minimal regions. In order to avoid false-positives due to hybridization “noise,” a minimum of two consecutive clones showing change was required for a region to be considered altered. This was made possible by the overlapping coverage of the tiling-path array.

Results and Discussion

Tiling-path array CGH characterizes genetic alterations. Complete SeeGH karyograms were generated for 24 microdissected archival specimens including 20 OSCCs and 4 matched normal connective tissues. Each karyogram summarizes copy number status for the 32,433 DNA segments providing a global view of these
tumor genomes. Many regions of alteration previously detected with conventional metaphase CGH were found in these profiles (9, 10). Both gross chromosomal aberrations as well as specific regional changes are apparent. Figure 1A presents the SeeGH karyogram of one tumor, 125T, with the remainder given as supplemental data online.6 In this profile, we observed alterations such as loss on 3p and 9p, two of the early genetic changes known to OSCC (11). The karyogram also revealed small and complex alterations. For example, chromosome 11 exhibits large regions of copy number loss on the long arm (11q), a gain encompassing 11q13.3 (Cyclin D1), and a small high copy number gain 11q22.2-22.3 (MMP cluster) as well as the breakpoint of a deletion at the telomere of 11p (Fig. 1B). The details seen in this example illustrate the value of the enhanced resolution provided by the tiling-path array.

To confirm the reliability of the whole genome array, we compared array CGH data from this study against those from our previous work. One tumor, 528T, previously analyzed using the 535 clone 3p-arm specific array (3), was compared with the 3p data generated by the whole genome array (with 957 clones covering 3p). Although only 104 clones were common between the two arrays, the two profiles are identical, illustrating the robustness of the array CGH approach (Fig. 1C).

Frequency plot analysis to identify recurring alterations. To facilitate the identification of regions of frequent gain or loss, a frequency plot of the genetic alterations at each of the 32,000 loci among the 20 OSCC samples was constructed (Fig. 2A). Generation of this frequency plot required first the use of aCGH-Smooth software to identify segmental imbalances in each genomic profile (see legend of Fig. 2), followed by frequency calculation and data display using SeeGH software version 2.2.2. In order to identify recurrent changes and to define minimal regions of alteration, frequently altered regions identified in the frequency plot were then confirmed in a multiple alignment analysis to verify the presence of these alterations in the original data.

Interestingly, the whole genome tiling-path array detected many more alterations in the OSCCs than previously reported using conventional techniques. For example, the total number of gains and losses detected with aCGH-Smooth in 125T was 312, consisting of 100 deletions and 212 gains. In contrast, reports in the literature for total number of alterations in oral tumors using conventional CGH have yielded values <20 (9, 10), whereas reports using interval marker array CGH have only focused on rare amplicons. The increased detection sensitivity in our study can be attributed to the whole genome tiling-path array. Since the array CGH data from this study against those from our previous work.

Defining anticipated alterations in OSCC. To show the sensitivity of the tiling array, we confirmed the presence of copy number changes known to be frequent in OSCC using SeeGH karyogram alignment. This allowed both the identification of minimal regions of alteration and the establishment of the precise boundaries of copy number change for each region.

Frequency plots yield values for total number of alterations in oral tumors using conventional techniques. For example, the total number of gains detected using the tiling-path array. First, a novel submegabase genetic alteration that differs in both the width and level of gain. The minimal region spans from clone RP11-16407 to RP11-535N12 and contains only five annotated genes including EGFR.

In another example, Fig. 2C shows three cases with gain at the Cyclin D1 locus at 11q13.3, again demonstrating a variance in both the width and level of copy number increase. This gain was frequent as shown by the peak at 11q13 in the frequency plot, present in 9 of 20 (45%) cases with high copy number amplifications in 7 cases. The 0.79 Mb amplicon defined by case 125T in Fig. 2C encompasses eight genes including Cyclin D1.

Multiple alignment of the tiling-path data at 7p11.2, which harbors EGFR in all 20 OSCCs allowed the identification of a 1.06 Mb minimal region of copy number gain present in 8 of 20 cases (40%), which is consistent with previous studies using array CGH (ref. 12; Fig. 2B). This shows a large improvement in the detection sensitivity and resolution compared with interval arrays which reportedly detected a low frequency of copy number gain present in just 4 of 89 tumors (5%; ref. 5). Figure 2B shows the alignment of three cases at 7p11.2 (161T, 793T, and 528T) illustrating segmental alteration that differs in both the width and level of gain. The minimal region spans from clone RP11-16407 to RP11-535N12 and contains only five annotated genes including EGFR.

Novel frequent regions of segmental alteration. In addition to these known regions, we also detected many frequent novel alterations. These alterations include copy number gains at 3q23, 5p15.2, 7p12.3-13, 7q21.2, and 7q35 and copy number losses at 2p15, 4q34.3, and 16q23.2. All regions were submegabase in size and were present in as few as 7, and as many as 15, of the 20 cases. These regions are listed in Supplementary Table S2 available online and may also represent alterations critical to oral tumorigenesis.

Figure 3A and B show two contrasting examples of copy number gains detected using the tiling-path array. First, a novel submegabase segmental gain was detected at 7q35 in 8 of 20 cases (40%). Samples 125T, 123T, and 486T each showed an amplified region containing a single gene, rho guanine nucleotide exchange factor 5
Figure 1. High-resolution SeeGH karyogram analysis of oral tumors. A, whole genome karyogram of tumor 12ST. SeeGH software was used to plot the log2 signal intensity ratios of tumor versus reference DNA at each BAC clone. Each dot represents the signal ratio for an individual BAC clone. The green and red lines are scale bars indicating log2 signal intensity ratios of +0.5 and −0.5, respectively. Multiple regions of copy number gain (5p and 7p) and loss (3p and 8p) are apparent, with the array revealing small and/or complex alterations. We also see gains at the hTERT (5p15.33) and cMYC (8q24.21) loci. B, small and complex alterations on chromosome 11 show the sensitivity of the array to detect detailed alterations, both high copy number and single copy number changes. Chromosome 11 exhibits multiple genetic alterations on both arms. Large regions of segmental single copy number loss and segmental gain are evident. Magnification of these segmental changes enables the identification of the precise location of breakpoints between neighboring BAC clones. Multiple focal changes such as a gain at 11q13.3 (containing CCND1) and a small novel high copy number gain 11q22.2-22.3 (MMP cluster) are also detected. C, comparison of 3p profiles of tumor 528T shows the robustness of the whole genome tiling-path array. These profiles were generated from the whole genome tiling-path array (3p data, 957 overlapping clones plotted according to base pair positions) in the current work (left) and the previously published 3p-arm-specific array, which consisted of 535 clones plotted in relative clone order (right, ref. 3). Although only 104 clones are common between these arrays, identical regions of alteration were identified. Examples are shaded green and red.
One case (486T) showed a high-level amplification that was absent in the matched non–cancer DNA derived from underlying connective tissue (Fig. 3A). ARHGEF5 is a member of the Dbl guanine nucleotide exchange factor family specific for Rho GTPases, which are well-known to be involved in tumorigenesis (14).

Second, and most noticeable, a 2.17-Mb-wide high copy gain at 11q22.2-22.3 was detected in two samples (125T and 805T) with a single copy number gain in an additional sample (628T; Fig. 3B). The extremely high copy number amplification at this region suggests its biological importance. A recent study on OSCC using a genome-wide interval marker array reported a rare amplicon (5.6% of cases) at this loci. However, here we report a much higher frequency (15%) and have fine-mapped the breakpoints in greater detail. In fact, the region we describe here contains 17 known genes, with two gene clusters, including nine matrix metalloproteinase (MMP) genes (MMP1, 3, 7, 8, 10, 12, 13, 20, and 27), and two baculoviral IAP repeat-containing protein (BIRC) genes (BIRC2 and 3). This is the first report of the coamplification of two

Figure 2. Frequency of gains and losses. A, summary of frequency of losses (green, left) and gains (red, right) detected by aCGH-Smooth software for each chromosome arm. The bars to the left and right of the chromosomes represent a frequency of 0.5 and 1, respectively. The aCGH-Smooth software identifies alterations and their associated breakpoints in the normalized data by employing a local search algorithm. By calculating the probability that a given clone is of the same copy number status as a set of previous clones, the algorithm calculates a maximum likelihood for each clone. This enables the identification of putative breakpoints which are then shifted randomly upstream and downstream to calculate an overall goodness of fit (8). Thus, groups of clones considered to have the same copy number status are assigned a relative value. Gains and losses for each clone were then tabulated for 20 OSCCs and plotted in SeeGH software (7) to generate a frequency plot for the identification of “hotspots” or regions of frequent gain and loss. B, alignment of three cases (161T, 793T, and 528T) showing gains (red) at 7p11.2 (EGFR locus) that differ in both width and level. The central bar and the red bars to the right and green bars to the left, represent log2 signal intensity ratios of 0, +0.5 and +1, and −0.5 and −1, respectively, between the sample and reference DNA. C, alignment of three cases (125T, 486T, and 542T) showing a gain at 11q13.3, CCND1 locus (red). D, alignment of losses at 8p23.2 in 5 cases (669T, 125T, 528T, 117T, 814T) with the deleted region shaded in green disrupting the 2.06 Mb CSMD1 gene.
gene clusters, both relevant to oral cancer. BIRC2 and BIRC3 are known to function in inhibition of apoptosis and have been shown to be amplified in lung cancer (15). MMP gene products function in invasion and metastasis through degradation of the basement membrane and extracellular matrix. MMP overexpression has previously been associated with oral cancer (16).

**Frequent amplification and overexpression of TRIO.** Multiple karyogram alignment of the 5p arm revealed a novel 0.58 Mb copy number gain at 5p15.2 present in 9 of 20 cases. Figure 4A shows the alignment of three such cases (125T, 161T, and 486T) showing the most narrow segmental gains detected at this region. The region spanning from RP11-744A15 to RP11-611H4, encompasses a single known gene, triple functional domain (TRIO), and is not seen in the matched normal samples. TRIO expression was significantly higher in the four OSCCs compared with the seven normal specimens (Fig. 4A). TRIO contains three functional domains—a serine/threonine kinase domain and two guanine exchange factor domains. TRIO amplification and abundant expression was recently reported in lung carcinoma and bladder cancer (17–19). Although 5p amplifications have been well-documented in oral cancer, this is the first report of the microamplification of the TRIO locus in oral cancer, probably due to insufficient resolution of both conventional techniques and interval marker array CGH.

**Frequent amplification and overexpression of CDK6.** Multiple karyogram alignment of the 7q arm revealed a 0.32 Mb copy number gain at 7q21.2 present in 6 of 20 cases. Although a “rare” 3 Mb amplicon has been previously described (5), we detected this region with much higher frequency (30%) and have fine-mapped precise breakpoints at RP11-82E23 and at RP11-514K1. The minimal region we describe encompasses two known genes, cyclin-dependent kinase 6 (CDK6) and peroxisome biogenesis factor 1 (PEX1). Figure 4B shows the alignment of three cases that exhibited the most narrow copy number gains (528T, 542T, and 486T). CDK6 has been shown to be overexpressed in numerous squamous cell carcinoma cell lines including OSCC cell lines (20), and is well known to positively regulate the cell cycle, emphasizing its potential role in oral tumorigenesis. In contrast, PEX1 functions in peroxisomal matrix protein import, which, if interrupted, results in peroxisomal biogenesis disorders and has not previously been implicated in tumorigenesis. Therefore, only CDK6 was further investigated. CDK6 expression was evaluated in four tumors and seven normal specimens—the same samples used in comparing TRIO expression. As shown, CDK6 expression was significantly higher in the tumors compared with the normal sample sets (Fig. 4B). Deregulation of TRIO and CDK6 at the transcriptional level shows the biological significance of these loci.

**Conclusion.** In this study, we presented the first application of whole genome tiling-path array CGH to clinical specimens. We interrogated 32,433 overlapping segments covering the genome for the detailed delineation of the genomic landscapes of 20 OSCCs.

![Figure 3](https://example.com/figure3.png)

**Figure 3.** Novel regions of segmental gain. A, alignment of 125T, 123T, and 486T, shows a novel submegabase gain at 7q35, unchanged in matched normal sample (486C) corresponding to 486T. Shaded red areas, copy number gain. Shaded green areas, losses; see legend of Fig. 2B for further explanation on ratio bars. The high level gain present in tumor 486T is absent in the matched normal. B, high level amplification at 11q22.2-22.3. Minimal region of alteration is marked by dashed lines and contains a cluster of 12 genes including 9 MMP genes and 2 BIRC genes (right).
This unprecedented resolution enabled analysis from the whole genome view, to individual chromosome arm views, and down to the gene level. At this resolution, we showed the presence of numerous genetic alterations common in OSCC too diminutive for detection by conventional cytogenetic techniques. The detection of submegabase segmental amplifications at 5p15.2, containing \textit{TRIO}\, at 7q21.2, containing \textit{CDK6}, and at 11q22.2-22.3, containing the \textit{MMP} gene cluster are prime examples. These results validate the need to employ tiling resolution analysis of tumor genomes in cancer gene-discovery.

\textbf{Acknowledgments}

Received 5/2/2005; revised 6/28/2005; accepted 7/6/2005.

Grant support: National Institute of Dental and Craniofacial Research grants R01 DE13124 and R01 DE015965; Canadian Institute of Health Research; Genome Canada/British Columbia; the Hardwick scholarship (C. Baldwin); the National Science Engineering Research Council (C. Garnis) and the Michael Smith Foundation for Health Research scholarships (C. Garnis).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked \textit{advertisement} in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank S. Watson, C. Malloff, R. DeLeeuw, B. Chi, P. Wang, J. Davies, and B. Coe for array synthesis and bioinformatics assistance.

\textbf{References}


Multiple Microalterations Detected at High Frequency in Oral Cancer

Corisande Baldwin, Cathie Garnis, Lewei Zhang, et al.


**Updated version** Access the most recent version of this article at:
[http://cancerres.aacrjournals.org/content/65/17/7561](http://cancerres.aacrjournals.org/content/65/17/7561)

**Supplementary Material** Access the most recent supplemental material at:
[http://cancerres.aacrjournals.org/content/suppl/2005/08/29/65.17.7561.DC1](http://cancerres.aacrjournals.org/content/suppl/2005/08/29/65.17.7561.DC1)

**Cited articles** This article cites 20 articles, 3 of which you can access for free at:
[http://cancerres.aacrjournals.org/content/65/17/7561.full#ref-list-1](http://cancerres.aacrjournals.org/content/65/17/7561.full#ref-list-1)

**Citing articles** This article has been cited by 15 HighWire-hosted articles. Access the articles at:
[http://cancerres.aacrjournals.org/content/65/17/7561.full#related-urls](http://cancerres.aacrjournals.org/content/65/17/7561.full#related-urls)

**E-mail alerts** Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions** To request permission to re-use all or part of this article, use this link [http://cancerres.aacrjournals.org/content/65/17/7561](http://cancerres.aacrjournals.org/content/65/17/7561). Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.