Human Molecular Genetics, 2014, Vol. 23, No. 10 **2618–2628**
doi:10.1093/hmg/ddt657
Advance Access published on January 8, 2014

Recurrent genomic alterations in sequential progressive leukoplakia and oral cancer: drivers of oral tumorigenesis?

Nilva K. Cervigne[1,4], Jerry Machado[5,8], Rashmi S. Goswami[5,9], Bekim Sadikovic[10], Grace Bradley[1,6], Bayardo Perez-Ordonez[2], Natalie Naranjo Galloni[11], Ralph Gilbert[12], Patrick Gullane[12], Jonathan C. Irish[12], Igor Jurisica[3,7], Patricia P. Reis[1,13] and Suzanne Kamel-Reid[1,2,5,*]

[1]Division of Applied Molecular Oncology, Ontario Cancer Institute, [2]Department of Pathology, Toronto General Hospital, Ontario Cancer Institute and [3]Princess Margaret Cancer Centre and Techna Institute, University Health Network, Toronto, ON, Canada, [4]Department of Oral Diagnostics, Faculty of Dentistry (FOP), UNICAMP, Piracicaba, SP, Brazil, [5]Department of Laboratory Medicine and Pathobiology, [6]Faculty of Dentistry and [7]Departments of Medical Biophysics and Computer Science, University of Toronto, ON, Canada, [8]Prevention Genetics, Marshfield, WI, USA, [9]Department of Hematopathology, Vancouver General Hospital, Vancouver, BC, Canada, [10]Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada, [11]Department of Otolaryngology, Hospital Calderon Guardia, San Jose, Costa Rica, [12]Department of Otolaryngology/Surgical Oncology, Princess Margaret Hospital, University of Toronto and University Health Network, Toronto, ON, Canada and [13]Department of Surgery and Orthopedics, Faculty of Medicine, São Paulo State University (UNESP), Botucatu, SP, Brazil

Received August 23, 2013; Revised November 26, 2013; Accepted December 20, 2013

A significant proportion (up to 62%) of oral squamous cell carcinomas (OSCCs) may arise from oral potential malignant lesions (OPMLs), such as leukoplakia. Patient outcomes may thus be improved through detection of lesions at a risk for malignant transformation, by identifying and categorizing genetic changes in sequential, progressive OPMLs. We conducted array comparative genomic hybridization analysis of 25 sequential, progressive OPMLs and same-site OSCCs from five patients. Recurrent DNA copy number gains were identified on 1p in 20/25 cases (80%) with minimal, high-level amplification regions on 1p35 and 1p36. Other regions of gains were frequently observed: 11q13.4 (68%), 9q34.13 (64%), 21q22.3 (60%), 6p21 and 6q25 (56%) and 10q24, 19q13.2, 22q12, 5q31.2, 7p13, 10q24 and 14q22 (48%). DNA losses were observed in >20% of samples and mainly detected on 5q31.2 (35%), 16p13.2 (30%), 9q33.1 and 9q33.29 (25%) and 17q11.2, 3p26.2, 18q21.1, 4q34.1 and 8p23.2 (20%). Such copy number alterations (CNAs) were mapped in all grades of dysplasia that progressed, and their corresponding OSCCs, in 70% of patients, indicating that these CNAs may be associated with disease progression. Amplified genes mapping within recurrent CNAs (*KHDRBS1*, *PARP1*, *RAB1A*, *HBEGF*, *PAIP2*, *BTBD7*) were selected for validation, by quantitative real-time PCR, in an independent set of 32 progressive leukoplakia, 32 OSSCs and 21 non-progressive leukoplakia samples. Amplification of *BTBD7*, *KHDRBS1*, *PARP1* and *RAB1A* was exclusively detected in progressive leukoplakia and corresponding OSCC. *BTBD7*, *KHDRBS1*, *PARP1* and *RAB1A* may be associated with OSCC progression. Protein–protein interaction networks were created to identify possible pathways associated with OSCC progression.

*To whom correspondence should be addressed at: 610 University Avenue, Rm. 9-622, Princess Margaret Hospital, Ontario Cancer Institute and University Health Network, Toronto, Ontario, Canada M5G 2M9. Tel: +1 4163404800; Fax: +1 4163403596; Email: suzanne.kamel-reid@uhn.ca

© *The Author 2014. Published by Oxford University Press.*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Head and neck squamous cell carcinomas (HNSCCs) are the sixth leading cause of cancer death worldwide (1,2). About one-fourth of all HNSCCs are oral squamous cell carcinomas (OSCCs), and are estimated >26 000 new cases and >5000 deaths in the USA every year (3) [Source: estimated new cases are based on 1995–2008 incidence rates from 47 states and the District of Columbia as reported by the North American Association of Central Cancer Registries (NAACCR), representing about 95% of the US population. Estimated deaths are based on US Mortality Data, 1994 to 2008, National Center for Health Statistics, Centers for Disease Control and Prevention.]. Patients with OSCC have benefited from the latest advances in surgical techniques, radiation therapy and chemotherapy, which help enhance quality of life and improve survival. Despite these advances, the 5-year survival rate of patients remains at ∼50% (4–6). Low survival rates are mainly due to the presence of late-stage disease at diagnosis and disease recurrence. In order to improve patient survival, more accurate methods of detection of lesions at a risk for malignant transformation and a better understanding of the genetic events associated with disease progression are needed. Since malignant transformation is due to genetic damage over time (7), the identification of genetic changes in sequential progressive lesions within the oral cavity is thus potentially useful for predicting lesions at a risk for malignant transformation.

It is known that a significant proportion (up to 62%) of OSCCs (8,9) arise from precursor oral potential malignant lesions (OPMLs), such as leukoplakia. Oral leukoplakia is a lesion that presents as a ‘white patch’ in the oral mucosa (9,10). Currently, these lesions are classified based on clinical and histopathological assessment. Clinically, leukoplakia lesions are homogeneous or non-homogeneous, the latter having a higher risk of transformation. Histologically, they are classified as non-dysplastic or dysplastic (10), and the presence of epithelial dysplasia is associated with an increased risk of malignant transformation of up to 31% (11). However, clinical and histological characteristics have limited prognostic value for predicting which leukoplakia will progress to malignancy.

Genetic biomarkers may be clinically useful to identify lesions at a risk for malignant transformation. Previous studies identified gains and losses of large chromosomal regions as well as loss of heterozygosity (LOH) associated with progression in dysplasias and OSCCs from different patients (12–14). As these studies identified fairly large genomic regions, important genes dominantly involved with OSCC progression remain unknown.

High-resolution global genomic profiling of sequential, progressive leukoplakia lesions and same-site invasive oral carcinoma from same patients enabled us to identify recurrent copy number gains and losses, and narrow down genes likely to be involved in progression of leukoplakia to invasive OSCC. Genes identified herein may represent prognostic markers, to identify leukoplakia lesions at a risk of progression. Assessment of such biomarkers could then be used to initiate early intervention, ultimately improving patient survival.

RESULTS

We successfully applied a whole genomic amplification (WGA) protocol for amplification of low yield DNA from formalin fixed paraffin-embedded (FFPE) samples, to accurately assess DNA copy number gains and losses. Array comparative genomic hybridization (aCGH) using amplified DNA allowed the identification of global copy number gains and losses, with similar results when compared with DNA from unamplified FFPE samples; results showed a high correlation between copy number alterations (CNAs) identified in amplified versus unamplified DNA samples ($R^2 = 0.80–0.97$).

CNAs were analyzed blinded to sample histology. Unsupervised hierarchical clustering analysis showed that the majority of progressive leukoplakia lesions (16/20) and OSCCs clustered together, separately from normal and non-progressive leukoplakia samples (Fig. 1), indicating that progressive leukoplakia and OSCCs share common CNAs.

The genomic segmentation algorithm used to detect amplifications and deletions showed a total of 8409 change calls in the group of progressive leukoplakia and OSCCs, and 2170 change calls in non-progressive samples. These results were then filtered for CNAs found in both groups that were very rare or absent among copy number variations (CNVs) found in the general population. This analysis showed that out of the 8409 change calls, 4081 (48.5%) were unique to progressive leukoplakia and same-site OSCCs; and that out of the 2170 change calls, 1146 were CNAs unique to non-progressive leukoplakia. These 1146 change calls present in non-progressive samples were then removed from the 4081 CNAs found within progressive leukoplakia and same-site OSCCs. Such an approach was used to accurately determine the genetic changes involved in oral cancer progression, since we selected CNAs specific to progressive leukoplakia and corresponding OSCCs and absent in non-progressive leukoplakia samples. This analysis showed a total of 2935 CNAs present in progressive leukoplakia and OSCCs, but not in non-progressive leukoplakia. A larger number of gains were common to progressive leukoplakia and OSCC (80%), in contrast to a small number of losses (20%) (Fig. 2). DNA losses were represented mainly on chromosomes 5q31.2 (35%), 16p13.2 (30%) and 9q33.1–9q33.2 (25%). The remaining DNA losses were found on chromosomes 17q11.2, 3p26.2, 18q21.1, 4q34.1 and 8p23.2 (20%) (Supplementary Material, Fig. S1A). DNA copy number gains were identified on chromosome 1p in 20/25 cases (80%) with high-level amplifications at 1p35 and 1p36. Amplifications were also found at 11q13.4 (68%), 9q34.13 (64%), 21q22.3 (60%) and 6p21 and 6q25 (56%). Other regions of chromosomal gain included 10q24, 19q13.2, 22q12, 5q31.2, 7p13 and 14q22 (48%) (Supplementary Material, Fig. S1B).

We observed an average of 113, 61, 153 and 178 significant change calls ($P < 0.001$) in sequential progressive samples of mild ($n = 4$), moderate ($n = 3$), severe ($n = 6$) leukoplakia and OSCCs ($n = 25$), across all patients. Figure 3 shows a representative example of CNAs found in the sequential progressive samples of patient 4. Samples containing foci of cells with two different grades (e.g. samples 4d, 10d and 10e) were included in the group of higher grade of dysplasia. Also, as severe oral dysplasia and carcinoma *in situ* have very similar histology and do not differ biologically, they were considered a unique group in our analysis.

We detected a total of 696 different chromosomal regions, commonly altered in progressive leukoplakia and OSCCs; 552/696 were regions of gain and 144/696 were regions of

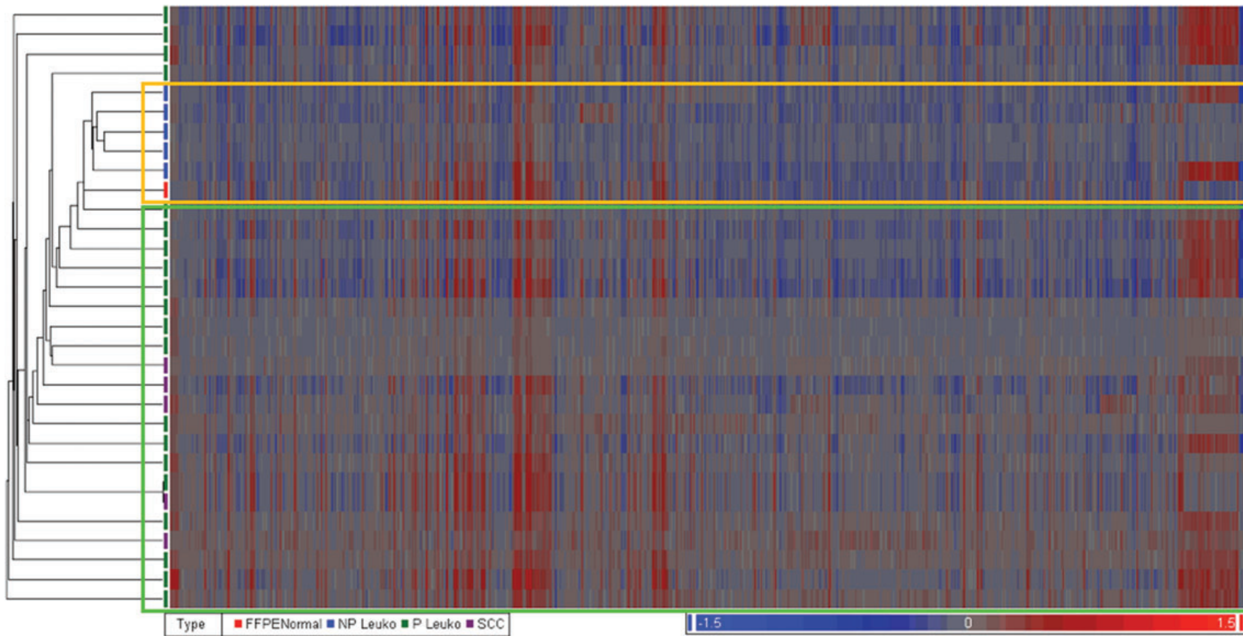

Figure 1. Unsupervised hierarchical clustering analysis of CNAs. Data were generated using Partek Genome Suite software. Two main distinct clusters were observed: normal/non-progressive leukoplakia (yellow box) and progressive leukoplakia/OSCCs (green box), demonstrating similarity between malignant and progressive lesions (purple and green, respectively), and showing how normal and non-progressive leukoplakia samples (red and blue, respectively) were clustered separately.

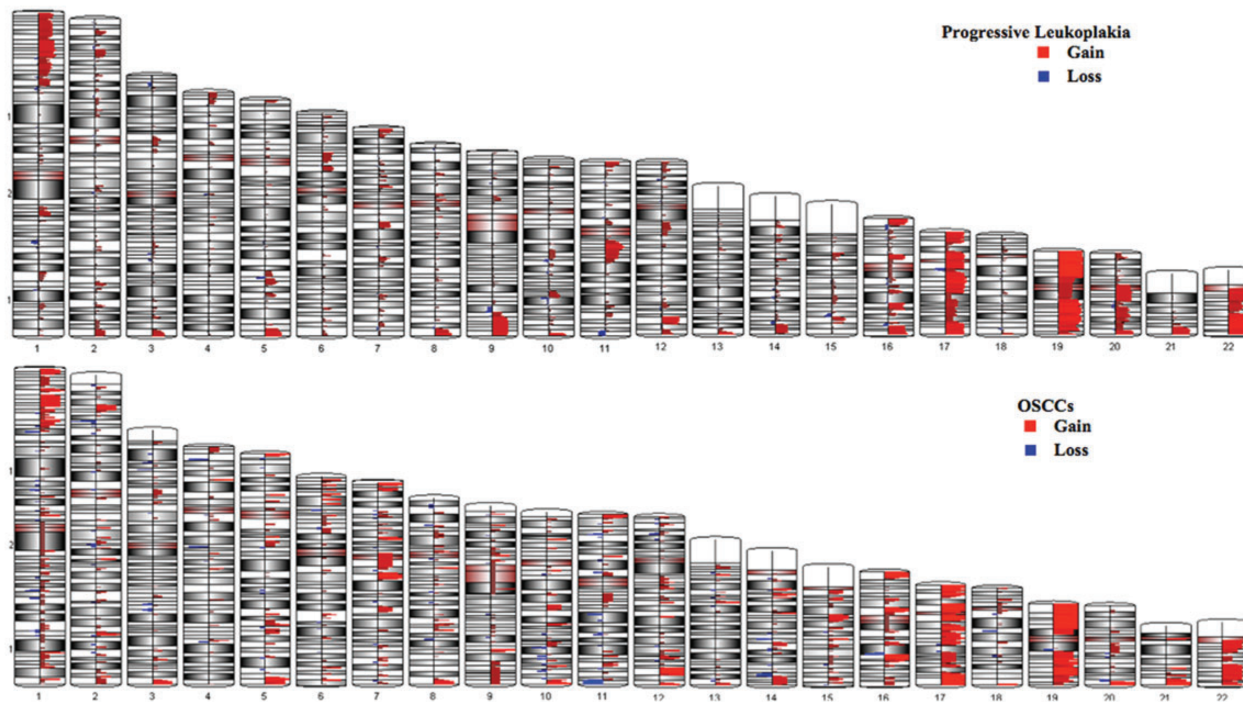

Figure 2. CNA profiles of progressive leukoplakia (upper panel) and OSCCs (bottom panel). CNA profiles were similar in progressive leukoplakia lesions and same-site OSCCs. Regions of gain (in red) are over-represented, compared with regions of loss (in blue).

loss. In order to determine the alterations involved in the progression of OPML to invasive carcinoma, we focused our analysis on CNAs from these 696 regions that were present in at least one OSCC and one preceding OPML from the same patient. This analysis revealed 193 regions of gains and 15 regions of loss. Notably, 38/193 gains and 5/15 losses were common to all sequential samples (OSCC and preceding leukoplakia lesions). This analysis confirmed recurrent losses mapped to 3p26.2, 8p23.2, 9q33.1–9q33.2, 17q11.2 and 18q21, and gains to 1q32, 1p35–36, 2p14, 5q31, 6p21, 6q25, 7p13, 10q24, 11q13.4, 12p13, 14q22, 19q13 and 22q12.3. These CNAs were detected in low to high grade dysplasias, and their corresponding

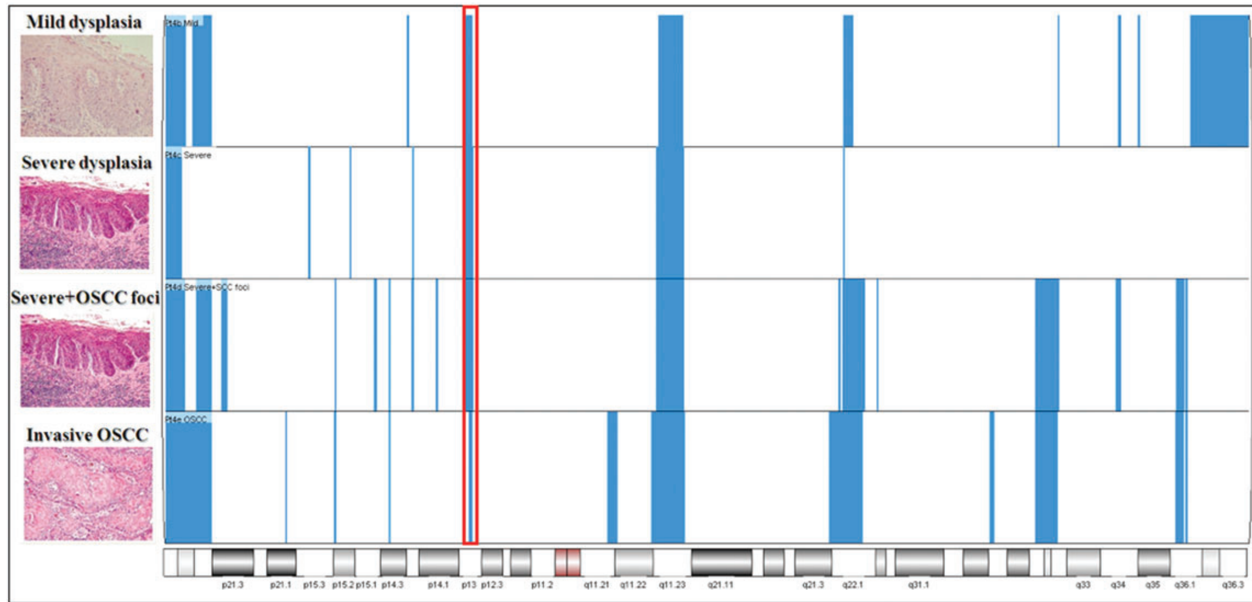

Figure 3. A representative example of CNA (7p13 gain) identified in sequential progressive samples of patient 4 (red box). 7p13 gains were detected in all progressive lesions from low to high grade sequential progressive leukoplakia and OSCCs, suggesting that 7p13 gains may be associated with disease progression. Other regions of 7p gains, detected in progressive samples from patient 4, were not highlighted as they were not present in all progressive leukoplakia from other patients.

OSCCs, for the majority (70%) of patients. Since these regions may contain genes that are relevant for the process of neoplastic transformation of leukoplakia to OSCC, we annotated 263 genes (255 amplified and 8 deleted), all mapped within the identified regions. We used the public databases UCSC Genome browser (http://genome.ucsc.edu/) and NCBI (http://www.ncbi.nlm.nih.gov/) as well as DAVID tool v6.7 (15,16) to search for and annotate the biological roles of these genes and their potential involvement in cancer biology. Based on these analyses, we selected 78 genes (Supplementary Material, Table S1), which were then subjected to further evaluation using ONCOMINE v.4 cancer profiling database (Research edition), a cancer microarray database and web-based data-mining platform aimed at facilitating discovery from genome-wide expression analyses (17). This analysis allowed us to verify whether deregulated mRNA expression of these 78 genes has been detected in independent samples from patients with head and neck cancer. Due to sample availability, given the rarity and small sample sizes of leukoplakia lesions, we were able to validate the top six genes from our analysis, which have also been described to play key roles in cancer-related pathways. Validation analysis was then performed in an independent cohort of progressive dysplasias as well as its corresponding OSCCs, and compared with a group of non-progressive leukoplakia lesions, using RQ-PCR (Fig. 4). *BTBD7, KHDRBS1*, *PARP1* and *RAB1A* were all found to be amplified in progressive leukoplakia lesions and OSSCs and not amplified in non-progressive leukoplakia. *NPM3* and *HBEGF* did not significantly differ between progressive and non-progressive leukoplakia.

A network-based analysis was used to map protein–protein interaction (PPI) networks of genes within CNAs commonly identified in progressive leukoplakia and OSCC. Such analysis may be useful to identify commonly disrupted genetic pathways in progressive leukoplakia and OSCC. PPI networks showed five proteins (HBEGF, KHDR1, PAIP2, PARP1 and RAB1A) interconnected through several common partner proteins; four of them shared the same partner protein (Supplementary Material, Fig. S2). Since we previously identified a 3-miRNA signature (hsa-miR-21, has-miR-345 and has-miR-281b) of oral leukoplakia progression using the same samples described herein for aCGH profiling (18), we sought to integrate PPI networks of genes and miRNAs. Using NAViGaTOR (19), we combined the network from Supplementary Material, Figure S2 with a network from proteins that correspond to targets of the 3-miRNA signature. This network-based integrative analysis identified regulatory networks that may be disrupted by CNAs during oral cancer progression. Our results showed that two miRNAs (hsa-miR-21 and hsa-miR-345) were highly interconnected and associated with three proteins (BTBD7, PARP1 and RAB1A) within the first network (Supplementary Material, Fig. S3). BTBD7 was linked to hsa-miR-345 and RAB1A was linked to hsa-miR-21, KHDR1 and PARP1, respectively. KHDR1, PARP1 and RAB1A were highly connected proteins within the PPI networks identified.

Tobacco consumption data in patients with progressive versus non-progressive leukoplakia

We performed an additional analysis to assess whether history of tobacco consumption was associated with progressive leukoplakia. We found that tobacco smokers were the significant majority of patients with progressive leukoplakia and same-site OSCC (89.4%, 17/19 patients), compared with patients with non-progressive leukoplakia ($P = 0.0051$, 95%CI, Fisher's exact test). Most patients with non-progressive leukoplakia did not have a history of tobacco consumption (61, 5%, 8/13).

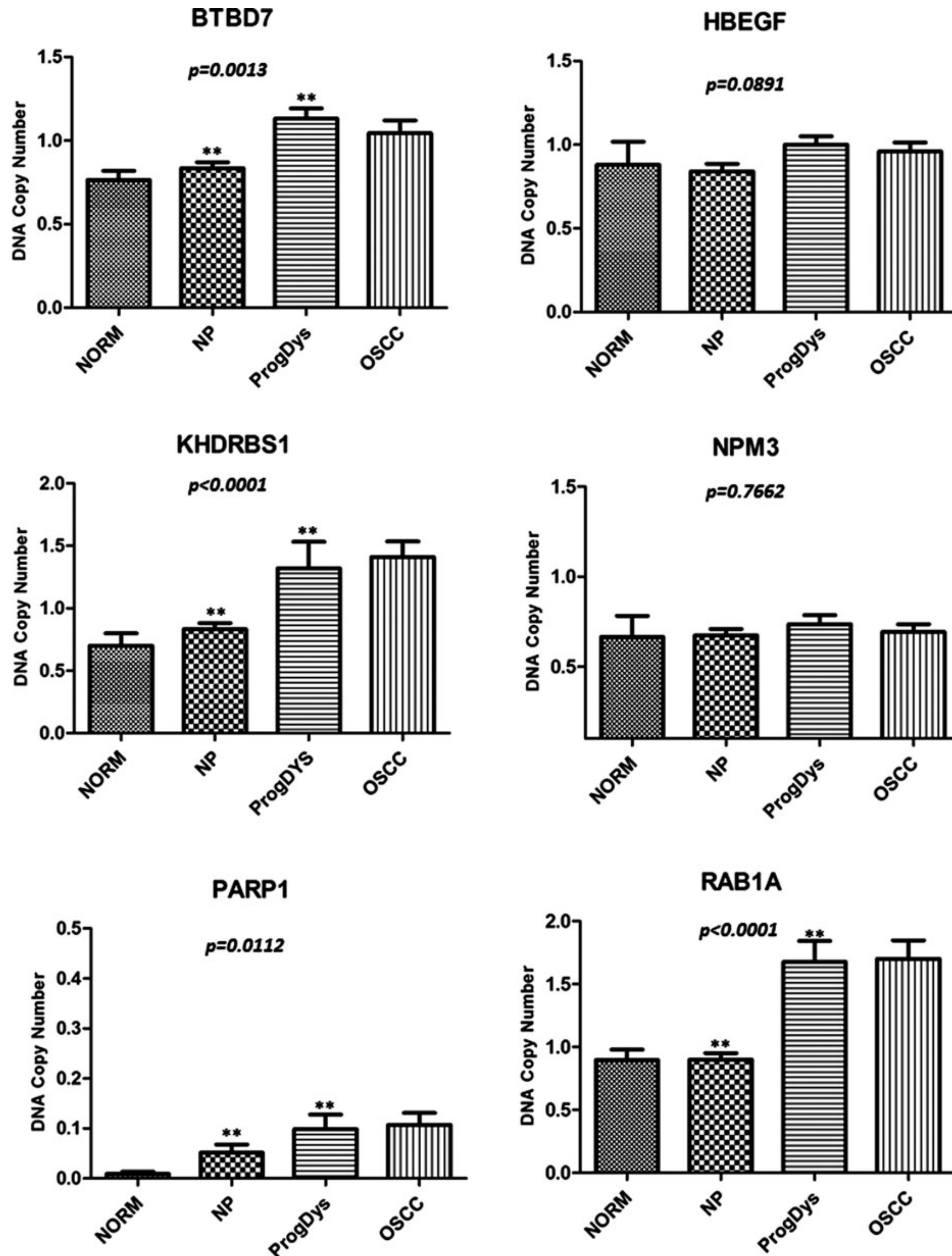

Figure 4. RQ-PCR validation of genes mapped within regions of gains, as identified by aCGH. DNA gains of *BTBD7*, *KHDRBS1*, *PARP1* and *RAB1A* were validated in an independent set of progressive leukoplakia and OSCCs compared with non-progressive leukoplakia samples. Asterisks indicate statistical significance for DNA copy number changes in progressive leukoplakia and OSCC compared with non-progressive leukoplakia and normal oral tissues (*P*-values are given by the Kruskal–Wallis test).

DISCUSSION

aCGH remains a high-throughput molecular technique widely used to determine CNAs across the cancer genome. This technology has been applied to DNA extracted from archived FFPE clinical specimens in several tumor types to elucidate key genes involved in disease development and progression (20,21). In particular, genomic analysis of tumor DNA has identified alterations in sequence and copy number associated with diagnosis, prognosis and treatment response in several cancers (22–24). Amplification or deletion of distinct chromosomal regions can lead to deregulated gene expression, thus conferring a growth advantage to malignant cells (25). Amplified or deleted genes could, therefore, be important targets for therapeutic intervention, and identification of such CNAs may also help elucidate potential mechanisms of tumor development and progression.

Our study is the first to examine CNAs in sequential, progressive oral lesions from the same patients, with the aim of identifying CNAs associated with malignant transformation, as the tissue progresses from benign epithelium to carcinoma. Non-progressive samples were also examined to ensure that CNAs were exclusively detected in progressive and not in non-progressive leukoplakia, and would thus likely represent driver changes associated with malignant transformation. Such changes provide targets for functional validation of molecular mechanisms of disease progression.

When we compared CNA profiles of histologically different tissues using unsupervised hierarchical clustering analysis, we were able to show that non-progressive leukoplakia segregated from progressive leukoplakia and invasive OSCCs. CNAs identified mainly increased from the lowest to the highest grade of dysplasia, culminating with invasive OSCCs, which contained the largest number of changes. Our data demonstrate that progressive leukoplakia lesions already possess many of the genetic alterations present in invasive cancers; this is consistent with the hypothesis that the majority of genetic alterations occur at an early phase during head and neck cancer progression (26). Leukoplakia lesions may thus have genetic instability, as it has been shown that chromosome instability is detected in the majority (9/10) of aneuploid OSCCs that arose from aneuploidy leukoplakia (27).

Tobacco smoking, an important risk factor associated with OSCC development, can lead to genomic instability and may help promote transformation of pre-malignant lesions in the oral cavity (28–30). Indeed, we showed that tobacco as well as alcohol consumption was associated with progressive lesions. Screening of patients with OPML who have history of tobacco and alcohol consumption may thus be important for early detection of oral cancer, ultimately improving patient survival (30).

Herein, we detected genomic instability with deletions on 3p, 9q and 18q in at least 20% of progressive leukoplakia and same-site OSCC. Allelic losses of 3p and 9p have been reported in HNSCC from different anatomic sites (31,32); 18q loss has been associated with poor prognosis and metastasis in patients with HNSCC (33,34). LOH on 9p, combined with *TP53* mutations, have been recently associated with malignant transformation of oral leukoplakia, and suggested as biomarkers of progression (35).

Over-representation of CNAs identified at lp35–36, 11q13, 19q and 22q12, in low-to-high-grade sequential progressive leukoplakia and OSCC, agrees with a previous study showing that these regions were correlated with HNSCC progression (31). In particular, 11q13 gains/amplification have been associated with poor prognosis of patients with HNSCC (36). Gains at 11q have been associated with a higher risk of esophageal squamous cell carcinoma development; notably, 11q13.1–13.4 is one of the most gene-rich regions on 11q, showing DNA gains correlated with increased RNA expression in $>80\%$ of the genes mapped on this minimal region (37). Additionally, 11q13 gains were correlated with poor prognosis of patients with other tumors, such as prostate (38) and thyroid (39), larger tumor size in hepatocellular carcinoma (40), and were suggested as predictive markers of distant recurrence in patients with breast cancer (41).

Our aCGH data analyses narrowed down 16 altered genes within regions containing CNAs associated with progressive leukoplakia and same-site invasive OSCC. Significantly amplified genes (*BTBD7*, *CAMSAP1L1*, *CHRDL2*, *FBXO7*, *GMPK2*, *HBEGF*, *IRF9*, *KHDRBS1*, *NPM3*, *PAIP2*, *PARP1*, *RAB1A*, *REC8* and *TBRG4*) and significantly deleted genes (*CSMD1* and *MYO5B*) were detected in progressive leukoplakia and OSCCs, and not in non-progressive lesions. Gene ontology (GO) categories for those genes include functions associated with protein binding, cell cycle, cell differentiation, cell proliferation, transcription factors and cell surface receptors linked to signal transduction (42) (Supplementary Material, Table S2). The most representative GO category comprised protein binding genes, which included *BTBD7*, *FBXO7*, *IRF9*, *PAIP2*, *PARP1*, *REC8* and *MYO5B*. The deleted gene, *CSMD1*, encodes a membrane–membrane interaction protein. *RAB1A* is involved in protein transportation and transduction signaling, and *KHDRS1, PARP1* and *IRF9* are transcription factor regulators. Additionally, *TBRG4* is involved in apoptosis; *CHRDL2* is a key player in cell differentiation, and *HBEGF* is part of the epidermal growth factor receptor/transduction signaling gene category.

Genetic alterations in sequential OPMLs and same-site OSCCs may represent key driver changes in disease development. In particular, DNA amplification can occur at an earlier stage of cancer development, in which oncogenic pathways may be critically disrupted (43). One of the most frequently (80%) amplified gene in our data set was *KHDRBS1* (also known as *p62*), a cell proliferation and cell surface receptor of signal transduction located at 1p35.1. The signaling adaptor p62 is induced by RAS, with p62 levels increased in human tumors, and required for RAS-induced survival and cellular transformation (44). Considering that the balance between cell death and survival is important in oncogenic transformation processes, assessment of *KHDRBS1-RAS* in oral cancer progression may be relevant to further elucidate the molecular mechanisms of oral tumorigenesis. Interestingly, *RAB1A*, a member of the RAS oncogene family, which maps to 2p14, was significantly amplified in progressive leukoplakia and OSCC from our data set. *RAB1A* overexpression has been previously identified in human tongue squamous cell cancer and suggested as a biomarker of tongue carcinogenesis (45).

Another amplified gene was *HBEGF*, mapped at 5q31.3. This gene encodes a protein that is an EGFR ligand up-regulated in diverse pathological conditions, including cancer (46,47). Interestingly, *HBEGF*, together with two other genes, *COX2* and *ST6GALNAC5*, mediate breast cancer cell passage through the blood-brain barrier (46). In addition, *COX-2* overexpression has been shown in oral cancer and high-risk oral lesions (48). Recently, a significant increase in COX-2 protein expression has been detected in moderate dysplastic oral leukoplakia when compared with inflammatory fibrous hyperplasia lesions and suggested to be associated with early stages of oral tumorigenesis (49).

EGFR overexpression is associated with poor prognosis of patients with HNSCC, and its inhibition improves patient survival (50). Increased EGFR-DNA copy number has been associated with OSCC development in patients having a precursor, same-site leukoplakia, which also over-expressed EGFR. EGFR inhibitors may thus be potentially useful in preventing malignant transformation of such leukoplakia lesions (51).

Additionally, *EGFR* is one of the most frequently amplified and highly expressed gene in both human and mouse oral tumors ($P < 0.01$) (52). Of note, *RAB1A* is a downstream effector of *EGFR*. Since amplification of *EGFR* and other oncogenes

may coexist in single cells of oral dysplasia (43), cells may thus be dependent on multiple oncogenes for OPML progression.

The roles of the *BTBD7, NPM3* and *PARP1* genes in tumorigenesis have also been described (53–55), although to our knowledge, *BTBD7* and *NPM3* have not been previously examined in oral tumorigenesis. *BTBD7* has been suggested to play a role in epithelial cell dynamics and branching morphogenesis, by inducing *SNAIL2* and suppressing E-cadherin expression; such events alter cell morphology and reduce cell-to-cell adhesion (56). Interestingly, podoplanin, a small mucin-type transmembrane protein that modulates molecular pathways of cell migration and invasion (57), is over-expressed in dysplastic oral leukoplakia and associated with grade of dysplasia and risk of progression to cancer (58).

Other cancer-associated genes identified in our aCGH experiments include *CHRDL2*, mapped on 11q13.4. *CHRDL2* encodes a putative extracellular matrix protein and was shown to be over-expressed in breast, lung and colon tumors, compared with corresponding normal tissues (59). *REC8*, located at 14q12, is part of the cohesin gene family. REC8 protein participates in the segregation of homologs at the first meiotic division (60), and *REC8* up-regulation determines the extent of arrested mitoses and polyploidy in lymphoma cell lines (61). *REC8*, together with *PAIP2*, are differentially expressed in OSCCs (62) and involved in cellular mitosis and translation-related activities. Translation initiation is regulated in response to mitogenic stimulation, and thus associated with cell cycle progression and cell growth. Combined over-expression of the cell-cycle related proteins TP53/p16(INK4a) as well as the proliferation marker Ki-67 was suggested as a marker of malignant transformation and able to classify high-risk leukoplakia (63). Expression changes in the components of the translational machinery can lead to global changes, such as an increase in protein expression and translational activation of mRNA and miRNA molecules that control cell growth and proliferation. Although translational control alterations occur in cancer, further investigation is required to determine their role in tumor development and progression (64).

Other studies have identified large chromosomal regions and LOH events in dysplasias and OSCCs, but did not characterize such changes in sequential, progressive samples from same patients (12–13,43,65). Additionally, CNA-associated genes have not been previously characterized into PPI networks and pathways. Herein, PPI network analysis showed that proteins encoded by *HBEGF*, *KHDRBS1*, *PAIP2*, *PARP1* and *RAB1A* are interconnected through several common partner proteins within a common network; most genes within this network have functions associated with genome maintenance, cellular fate and organization and transcriptional control.

In the previous study (18), we identified a 3-miRNA signature associated with the progression of oral leukoplakia to same-site OSCC. We showed that increased expression levels of hsa-miR-21, hsa-miR-181b and hsa-miR-345 were significantly associated with increased lesion severity during progression of leukoplakia to OSCC. Since changes in miRNA expression may occur through several mechanisms, such as transcriptional or post-transcriptional regulation, and changes in the expression of miRNA biogenesis enzymes; such mechanisms may be partially attributed to genomic gains or losses. Herein, aCGH global genomic approach allowed us to integrate CNAs associated with sequential progressive lesions from different patients, with the previously identified 3-microRNA signature (miR-21, miR-181b and miR-345) associated with leukoplakia progression, in the same samples. Interestingly, both miR-345 (14q32.2) and miR-181b (1q32.1) map to regions of DNA gain identified by aCGH in progressive leukoplakia and OSCC, suggesting that over-expression of these two miRNAs may be due to gains at these genomic loci. CNAs were absent in the chromosomal region where miR-21 maps, suggesting that in progressive leukoplakia and OSCC, miR-21 could be regulated by other transcriptional and/or post-transcriptional changes. A network-based integrative analysis can be used to identify pathways that may be disrupted by CNAs, including miRNA-mRNA targets, during disease progression. Using PPIs, our study is highly suggestive of a connection between specific CNAs, miRNAs and gene expression, thus shedding light on our understanding of leukoplakia to OSCC progression. Recently, higher expression of miR-21, miR-181b or miR-345 in leukoplakia has been associated with cytological features used to grade leukoplakia; e.g. increased number of mitotic figures, increased nuclear to cytoplasmic ratio or hyperchromasia, nucleoli number and size (66). Interestingly, miR-181 over-expression has been reported to enhance lymph-node metastasis in OSCC through cell migration (67).

A consistent pattern of changes on selected chromosome arms enabled us to identify specific genes involved in the OSCC progression. Potential genomic markers of interest were identified on chromosomes 1p, 2p, 5q, 8p, 11q, 14q, 18q and 22q and may represent drivers involved in oral cancer progression. By integrating data on CNAs and the previously identified 3-miRNA expression signature, we were able to show a possible interaction between these changes; these may represent alterations to genomic and post-transcriptional control mechanisms of specific genes and pathways important for disease progression. Functional studies directly targeting these genes/pathways will help clarify their role in the progression of leukoplakia to OSCC.

MATERIALS AND METHODS

Patient samples

Training sample set

All patient samples were FFPE tissues. We collected 30 samples from 10 patients. Of these, 25 were sequential samples from 5 patients (20 progressive leukoplakia and 5 same-site carcinomas); therefore, all carcinomas had at least one corresponding premalignant oral leukoplakia. Of the 20 leukoplakia lesions, 4 were non-dysplastic and 16 were dysplastic (mild, moderate or severe). The remaining five samples were non-progressive leukoplakia lesions, from five patients. Training sample set characteristics are detailed in Table 1. A commercial normal genomic DNA (Promega, Madison, WI, USA) was used as control in the aCGH experiments.

DNA isolation from FFPE samples

All samples underwent histopathological analysis by an experienced oral pathologist (G.B.) to ensure the presence of dysplasia or carcinoma in at least 80% of each tissue section. Samples were needle macro-dissected to select the target cell population for DNA extraction and genomic analysis. In short, genomic DNA

Table 1. Training sample set characteristics: sequential oral leukoplakia and same-site OSCCs

Patient	Sample ID	Site	Histopathological diagnosis	Group	Date of biopsy (month/year)	Age	Gender	Tobacco
1	1a	Tonsil	Focal keratosis[a]	PL	10.2001	83	F	Yes
	1b	Alveolus + FOM	Mild squamous hyperplasia[a]	PL	07.2003			
	1c	Anterior FOM	Carcinoma *in situ*	PL	05.2004			
	1d	Anterior FOM	Invasive moderately differentiated OSCC	OSCC	05.2004			
4	4a	Tongue	Severe squamous dysplasia	PL	11.1997	40	M	No
	4b	Tongue	Mild squamous dysplasia	PL	10.1998			
	4c	Tongue	Severe squamous dysplasia	PL	11.1998			
	4d	Tongue	Severe dysplasia + SCC micro foci	PL	11.1998			
	4e	Tongue	Invasive moderately differentiated OSCC	OSCC	02.2000			
5	5a	Right lateral tongue	Keratosis[a]	PL	03.1993	65	F	Yes
	5b	Right lateral tongue	Moderate dysplasia	PL	03.1993			
	5c	FOM	Severe dysplasia	PL	10.1994			
	5d	Buccal mucosa	Mild dysplasia	PL	02.1997			
	5e	Left buccal mucosa	Invasive OSCC	OSCC	10.1997			
10	10a	Left buccal mucosa	Keratosis[a]	PL	03.1991	60	M	Yes
	10b	Left buccal mucosa	Moderate dysplasia	PL	09.1995			
	10c	Left buccal mucosa	Mild dysplasia	PL	09.1996			
	10d	Left buccal mucosa	Moderate + Severe dysplasia	PL	10.2003			
	10e	Left buccal mucosa	Severe dysplasia + carcinoma *in situ*	PL	11.2003			
	10f	Left buccal mucosa	Invasive moderately differentiated OSCC	OSCC	06.2004			
15	15a	Right tongue	Mild dysplasia	PL	04.1994	82	F	Yes
	15b	Right tongue	Moderate dysplasia	PL	04.1994			
	15c	Right tongue	Keratosis mild atypia	PL	02.2001			
	15d	Right tongue	Moderately differentiated OSCC	PL	10.2003			
	15e	Right tongue	Mild squamous hyperplasia[a]	OSCC	03.2004			
1NP		Mandible gingiva	Mild dysplasia	NPL	1997	32	M	No
2NP		Mandible lingual mucosa	Mild dysplasia	NPL	2000	58	M	No
3NP		Buccal mucosa	Moderate dysplasia	NPL	2001	49	M	No
4NP		FOM	Moderate dysplasia	NPL	2001	60	M	No
5NP		FOM	Severe dysplasia	NPL	2001	63	F	No

PL, progressive leukoplakia; NPL, non-progressive leukoplakia; OSCC, oral squamous cell carcinoma; FOM, floor of mouth; WD, well differentiated; MD, moderately differentiated; PD, poorly differentiated; F, female; M, male.
[a]Non-dysplastic lesions.

was isolated from 5 to 10 (10 µm thick) FFPE tissue sections. After xylene deparaffinization, tissues were incubated in Cell Lysis Solution buffer (5 PRIME, Gaithersburg, MD, USA) and Proteinase K solution (20 mg/ml) for 2 days at 56°C (fresh aliquots of proteinase K were added at 17 and 24 h). Genomic DNA was isolated and purified using the ArchivePure DNA Cell/Tissue Kit-4g (5 PRIME), with final elution into sterile ddH_2O. DNA samples were quantified using a NanoDrop Spectrophotometer, and checked by agarose gel electrophoresis for quality. All samples yielded DNA of sufficient quantity and quality for analysis. Samples within the training set were subjected to WGA using a Sequenase-based approach (modified from the Affymetrix Chromatin Immunoprecipitation Assay, as per Sadikovic *et al*. (68), in order to yield enough DNA quantity for aCGH analysis (~2 µg). High-quality normal male genomic DNA (Promega) was used as the reference sample, as described in other aCGH studies (20,69,70). Male genomic DNA (Promega) was heat fragmented for 10 min at 95°C, subsequently subjected to WGA and hybridized against each test sample.

Sequenase-based WGA

Two rounds of WGA were used to randomly amplify the 30 FFPE DNA samples from the training set (minimum of 10 ng and maximum of 200 ng of DNA). This amplification protocol has been successfully used by others to amplify $<$10 ng of DNA and was utilized in the comparison of relative enrichment between two samples (71). The protocol comprises of two sets of enzymatic reactions (Supplementary Material, Table S3); in Round I, the sequenase enzyme is used to extend randomly annealed primers (Primer A) and to generate templates for subsequent PCR. During Round II, the specific primer B (the sequence of which is partially identical to Primer A) was used to amplify the templates previously generated by dNTP (10 mM) incorporation. Following each amplification round, DNA was purified using the QIAquick® PCR Purification Kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol. The final purified PCR product was eluted into 50 µl of Sigma water, and 5 µl of product was used to run a 1% agarose gel, to verify the presence of a 200 bp–1 kb DNA 'smear' for successfully amplified samples.

To verify the fidelity of the WGA, we first sought to determine the correlation of data resulting from amplified and unamplified template DNA, by analyzing matched fresh-frozen and FFPE tumor samples from a same patient. In all experiments, WGA protocol was used for both test (tumor) and reference (Promega DNA) samples. Correlation data are given in Supplementary Material, Table S4. All WGA samples displayed adequate signal-to-background ratios. Array CGH profiles of paired samples did not display any chromosomal gains or losses due to WGA. Overall, we obtained consistent data when

comparing amplified versus unamplified FFPE samples, resulting in high Pearson's correlation coefficient ranging between $R^2 = 0.80–0.97$ for both samples (29T and 201T). These values were reflected in the mean absolute deviations of the log2 ratios, calculated for all probes across the genome on the array.

Whole genome tiling array CGH

We used the NimbleGen 385K whole genome tiling v2.0 array, which contains over 385 000 oligonucleotides probes (60-mer with a median probe spacing of ~7 kb) providing genome wide coverage. Array CGH experiments including quality control, DNA labeling, hybridization, scanning and data extraction were performed by NimbleGen Systems core facility (Reykjavik, Iceland). The complete experimental protocol is provided in the NimbleGen Arrays User's guide (https://projects.cgb.indiana.edu/download/attachments/5363/NimbleGen_CGH_Users_Guide_v3p1.pdf?version=2). Briefly, 1 μg of genomic DNA was used for dual color labeling (inverse Cy3/Cy5). All 30 samples were successfully labeled, meeting quality control criteria. Following hybridization, washing and scanning were performed according to the manufacturer's protocol (NimbleGen-Roche). Array CGH data generation was performed using commercially available software (SignalMap version 1.8, Nimblegen).

aCGH copy number data analysis

Partek Genomics Suite (PGS) software was used to identify CNAs. First, the *.pair* data files were loaded into the PGS software, which automatically loaded log2 ratio intensities for all probes across the tiling array. We performed unsupervised hierarchical clustering analysis using Euclidean distance, average linkage, agglomerative method (PGS), blinded to sample identity. We first sought to identify CNAs associated with oral cancer progression. For this, we performed copy number analysis across all samples: progressive leukoplakia lesions with corresponding OSCCs ($n = 25$), and non-progressive leukoplakia samples ($n = 5$). We examined CNAs present in progressive leukoplakia and corresponding OSCCs, and absent in non-progressive leukoplakia. We then compared these data against CNV frequency data available from the general control population [2115 predominantly European background individuals; half from Ontario (19), and half from Germany (20)]. This analysis filtered out any CNVs present in the general population, which are not relevant to disease biology/tumorigenesis. Additionally, to map the genetic alterations occurring during progression, we assessed CNAs within the progressive samples from each patient.

CNAs were identified using the genomic segmentation algorithm in PGS. Genomic aberrations were assessed with a segmentation stringency of 10 consecutive genomic markers utilizing $P < 0.001$ as cut off, and a signal-to-noise ratio cut-off of 0.3 for amplifications and deletions. We used a copy number cut-off of two copies to identify gains and losses; ratios < 0.85 are considered regions of loss, whereas ratios > 1.15 represent regions of gain. This analysis excluded genes mapped on sex chromosomes, and regions with no known genes.

Validation of genes by real-time quantitative PCR (RQ-PCR)

An independent cohort of 49 patient samples (21 non-progressive leukoplakia, 28 paired progressive leukoplakia lesions and OSCCs) was used for validation of CNAs identified by aCGH. Patient samples characteristics are described in Supplementary Material, Table S5. DNA was isolated from FFPE samples as discussed previously. We annotated and identified 16 altered genes, which were represented by several probes (at least 10) within the regions of recurrent gains and losses. Of these, 14 genes, including *BTBD7*, *CAMSAP1L1*, *CHRDL2*, *GMPK2*, *FBXO7*, *HBEGF*, *IRF9*, *KHDRBS1*, *NPM3, PAIP2*, *PARP1*, *RAB1A*, *REC8* and *TBRG4* were amplified, and 2 genes *CSMD1* and *MYO5B* were deleted in progressive leukoplakia lesions and paired OSCCs, compared with Promega normal gDNA, but not in non-progressive samples (Supplementary Material, Table S6). We further selected 6 of these 14 genes (*BTBD7*, *HBEGF*, *KHDRBS1*, *PAIP2*, *PARP1* and *RAB1A*), which were mapped within CNA regions exclusively amplified in all sequential progressive leukoplakia lesions and OSCCs and not amplified in non-progressive leukoplakia samples, for validation by RQ-PCR. RQ-PCR was performed using TaqMan assays (Life Technologies, Foster City, CA, USA); the detailed protocol is shown in Supplementary Material, Table S7. Primer/probe sequences are available upon request. Amplification conditions were: 50°C for 2 min.; 95°C for 10 min.; 35 cycles at 95°C for 15 s followed by 60°C for 1 min. RQ-PCR was performed using the Applied Biosystems Gene Amp PCR System 9700 thermocycler. Analysis was performed using the Delta-Delta Ct method (72). Statistical analysis was performed in GraphPad Prism v5.01 using the non-parametric Kruskal–Wallis test and Fisher's exact test. The significance level was $P < 0.05$. DNA copy number was normalized to Promega DNA control.

Protein–protein interaction network and pathways analyses

Genes selected for validation (*BTBD7*, *HBEGF*, *KHDRBS1*, *PAIP2*, *PARP1* and *RAB1A*) were mapped into corresponding proteins to generate PPI networks using I^{2}D ver. 2.0 (http://ophid.utoronto.ca/i2d) (73,74) and visualized using NAViGaTOR ver. 2.14 (http://ophid.utoronto.ca/navigator) (19). We further integrated three previously identified miRNAs (hsa-miR-21, hsa-miR-181, hsa-345) (18) into PPI networks, in order to identify whether genes and miRNAs shared common regulatory networks. This was achieved by first mapping miRNAs to target genes using microRNA Data Integration Portal ver.1 (http://ophid.utoronto.ca/mirDIP) (75), and then mapping resulting genes to proteins and interaction networks in I2D. We integrated predicted mRNA targets using mirDIP with the aim of identifying other potential partner proteins into regulatory networks, which may contribute to leukoplakia progression. In addition, The Database of Annotation, Visualization and Integrated Discovery (DAVID) v6.7 was used for functional annotation of genes (15,16).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Galloway Fund administered through the University Health Network (R.G. and S.K.R.), by the Cancer Research Society (Canada) (S.K.R.), and by Ontario Research Fund (GL2-01-030) (I.J.). Computational resources were supported in part by grants from the Canada Foundation for Innovation (#12301, #203373, #29272, #225404) and IBM (I.J.). Funding to pay the Open Access publication charges for this article was provided by OICR (Ontraio institute for cancer research), The Galloway Research Fund PMH Foundation, The Cancer Research Society and The Ontario MOHLTC.

REFERENCES

1. Parkin, D.M. (2001) Global cancer statistics in the year 2000. *Lancet Oncol.*, **2**, 533–543.
2. Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E. and Forman, D. (2011) Global cancer statistics. *CA Cancer J. Clin.*, **61**, 69–90.
3. Carvalho, A.L., Nishimoto, I.N., Califano, J.A. and Kowalski, L.P. (2005) Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database. *Int. J. Cancer*, **114**, 806–816.
4. Atula, T., Silvoniemi, P., Kurki, T., Varpula, M. and Grénman, R. (1997) The evaluation and treatment of the neck in carcinoma of the oral cavity. *Acta Otolaryngol. Suppl.*, **529**, 223–225.
5. Brunin, F., Mosseri, V., Jaulerry, C., Point, D., Cosset, J.M. and Rodriguez, J. (1999) Cancer of the base of the tongue: past and future. *Head Neck*, **21**, 751–759.
6. Forastiere, A., Koch, W., Trotti, A. and Sidransky, D. (2001) Head and neck cancer. *N. Engl. J. Med.*, **345**, 1890–1900.
7. Silverman, S. and Sugerman, P.B. (2000) Oral premalignancies and squamous cell carcinoma. *Clin. Dermatol.*, **18**, 563–568.
8. Lee, J.J., Hong, W.K., Hittelman, W.N., Mao, L., Lotan, R., Shin, D.M., Benner, S.E., Xu, X.C., Lee, J.S., Papadimitrakopoulou, V.M. *et al.* (2000) Predicting cancer development in oral leukoplakia: ten years of translational research. *Clin. Cancer Res.*, **6**, 1702–1710.
9. Schepman, K., der Meij, E., Smeele, L. and der Waal, I. (1999) Concomitant leukoplakia in patients with oral squamous cell carcinoma. *Oral Dis.*, **5**, 206–209.
10. Reibel, J. (2003) Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics. *Crit. Rev. Oral Biol. Med.*, **14**, 47–62.
11. Mithani, S.K., Mydlarz, W.K., Grumbine, F.L., Smith, I.M. and Califano, J.A. (2007) Molecular genetics of premalignant oral lesions. *Oral Dis.*, **13**, 126–133.
12. Tsui, I.F.L., Rosin, M.P., Zhang, L., Ng, R.T. and Lam, W.L. (2008) Multiple aberrations of chromosome 3p detected in oral premalignant lesions. *Cancer Prev. Res. (Phila)*, **1**, 424–429.
13. Garnis, C., Chari, R., Buys, T.P.H., Zhang, L., Ng, R.T., Rosin, M.P. and Lam, W.L. (2009) Genomic imbalances in precancerous tissues signal oral cancer risk. *Mol. Cancer*, **8**, 50.
14. Smith, J., Rattay, T., McConkey, C., Helliwell, T. and Mehanna, H. (2009) Biomarkers in dysplasia of the oral cavity: a systematic review. *Oral Oncol.*, **45**, 647–653.
15. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
16. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
17. Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A. and Chinnaiyan, A.M. (2004) ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*, **6**, 1–6.
18. Cervigne, N.K., Reis, P.P., Machado, J., Sadikovic, B., Bradley, G., Galloni, N.N., Pintilie, M., Jurisica, I., Perez-Ordonez, B., Gilbert, R. *et al.* (2009) Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma. *Hum. Mol. Genet.*, **18**, 4818–4829.
19. Brown, K.R., Otasek, D., Ali, M., McGuffin, M.J., Xie, W., Devani, B., Toch, I.L., van Toch, IL and Jurisica, I. (2009) NAViGaTOR: Network Analysis, Visualization and Graphing Toronto. *Bioinformatics*, **25**, 3327–3329.
20. Paris, P.L., Sridharan, S., Scheffer, A., Tsalenko, A., Bruhn, L. and Collins, C. (2007) High resolution oligonucleotide CGH using DNA from archived prostate tissue. *Prostate*, **67**, 1447–1455.
21. Johnson, N.A., Hamoudi, R.A., Ichimura, K., Liu, L., Pearson, D.M., Collins, V.P. and Du, M.-Q. (2006) Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells. *Lab. Invest.*, **86**, 968–978.
22. Albertson, D.G. and Pinkel, D. (2003) Genomic microarrays in human genetic disease and cancer. *Hum. Mol. Genet.*, **12**(Spec No.), R145–R152.
23. Forozan, F., Mahlamäki, E.H., Monni, O., Chen, Y., Veldman, R., Jiang, Y., Gooden, G.C., Ethier, S.P., Kallioniemi, A. and Kallioniemi, O.P. (2000) Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data. *Cancer Res.*, **60**, 4519–4525.
24. Lucito, R., Healy, J., Alexander, J., Reiner, A., Esposito, D., Chi, M., Rodgers, L., Brady, A., Sebat, J., Troge, J. *et al.* (2003) Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res.*, **13**, 2291–2305.
25. Schwab, M. (1990) Amplification of the MYCN oncogene and deletion of putative tumour suppressor gene in human neuroblastomas. *Brain Pathol.*, **1**, 41–46.
26. Ha, P.K., Benoit, N.E., Yochem, R., Sciubba, J., Zahurak, M., Sidransky, D., Pevsner, J., Westra, W.H. and Califano, J. (2003) A transcriptional progression model for head and neck cancer. *Clin. Cancer Res.*, **9**, 3058–3064.
27. Siebers, T.J.H., Bergshoeff, V.E., Otte-Höller, I., Kremer, B., Speel, E.J.M., van der Laak, J.A.W.M., Merkx, M.A.W. and Slootweg, P.J. (2013) Chromosome instability predicts the progression of premalignant oral lesions. *Oral Oncol.*, **49**, 1121–1128.
28. Saraswathi, T.R., Ranganathan, K., Shanmugam, S., Sowmya, R., Narasimhan, P.D. and Gunaseelan, R. (2006) Prevalence of oral lesions in relation to habits: cross-sectional study in South India. *Indian J. Dent. Res.*, **17**, 121–125.
29. Ahmadi-Motamayel, F., Falsafi, P., Hayati, Z., Rezaei, F. and Poorolajal, J. (2013) Prevalence of oral mucosal lesions in male smokers and nonsmokers. *Chonnam Med. J.*, **49**, 65–68.
30. Sujatha, D., Hebbar, P.B. and Pai, A. (2012) Prevalence and correlation of oral lesions among tobacco smokers, tobacco chewers, areca nut and alcohol users. *Asian Pac. J. Cancer Prev.*, **13**, 1633–1637.
31. Bockmühl, U., Schwendel, A., Dietel, M. and Petersen, I. (1996) Distinct patterns of chromosomal alterations in high- and low-grade head and neck squamous cell carcinomas. *Cancer Res.*, **56**, 5325–5329.
32. Califano, J., van der Riet, P., Westra, W., Nawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W. *et al.* (1996) Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res.*, **56**, 2488–2492.
33. Soder, A.I., Hopman, A.H., Ramaekers, F.C., Conradt, C. and Bosch, F.X. (1995) Distinct nonrandom patterns of chromosomal aberrations in the progression of squamous cell carcinomas of the head and neck. *Cancer Res.*, **55**, 5030–5037.
34. Van Dyke, D.L., Worsham, M.J., Benninger, M.S., Krause, C.J., Baker, S.R., Wolf, G.T., Drumheller, T., Tilley, B.C. and Carey, T.E. (1994) Recurrent cytogenetic abnormalities in squamous cell carcinomas of the head and neck region. *Genes Chromosomes Cancer*, **9**, 192–206.
35. Graveland, A.P., Bremmer, J.F., de Maaker, M., Brink, A., Cobussen, P., Zwart, M., Braakhuis, B.J.M., Bloemena, E., van der Waal, I., Leemans, C.R. *et al.* (2013) Molecular screening of oral precancer. *Oral Oncol.*, **49**, 1129–1135.
36. Akervall, J.A., Jin, Y., Wennerberg, J.P., Zätterström, U.K., Kjellén, E., Mertens, F., Willén, R., Mandahl, N., Heim, S. and Mitelman, F. (1995) Chromosomal abnormalities involving 11q13 are associated with poor prognosis in patients with squamous cell carcinoma of the head and neck. *Cancer*, **76**, 853–859.
37. Hu, N., Wang, C., Ng, D., Clifford, R., Yang, H.H., Tang, Z.-Z., Wang, Q.-H., Han, X.-Y., Giffen, C., Goldstein, A.M. *et al.* (2009) Genomic

characterization of esophageal squamous cell carcinoma from a high-risk population in China. *Cancer Res.*, **69**, 5908–5917.
38. Paris, P.L., Sridharan, S., Hittelman, A.B., Kobayashi, Y., Perner, S., Huang, G., Simko, J., Carroll, P., Rubin, M.A. and Collins, C. (2009) An oncogenic role for the multiple endocrine neoplasia type 1 gene in prostate cancer. *Prostate Cancer Prostatic Dis.*, **12**, 184–191.
39. Lee, J.-J., Au, A.Y.M., Foukakis, T., Barbaro, M., Kiss, N., Clifton-Bligh, R., Staaf, J., Borg, A., Delbridge, L., Robinson, B.G. *et al.* (2008) Array-CGH identifies cyclin D1 and UBCH10 amplicons in anaplastic thyroid carcinoma. *Endocr. Relat. Cancer*, **15**, 801–815.
40. Tanaka, Y., Kanai, F., Tada, M., Tateishi, R., Sanada, M., Nannya, Y., Ohta, M., Asaoka, Y., Seto, M., Shiina, S. *et al.* (2008) Gain of GRHL2 is associated with early recurrence of hepatocellular carcinoma. *J. Hepatol.*, **49**, 746–757.
41. Hwang, K.-T., Han, W., Cho, J., Lee, J.W., Ko, E., Kim, E.K., Jung, S.-Y., Jeong, E.-M., Bae, J.-Y., Kang, J.J. *et al.* (2008) Genomic copy number alterations as predictive markers of systemic recurrence in breast cancer. *Int. J. Cancer*, **123**, 1807–1815.
42. Higgins, M.E., Claremont, M., Major, J.E., Sander, C. and Lash, A.E. (2007) CancerGenes: a gene selection resource for cancer genome projects. *Nucleic Acids Res.*, **35**, D721–D726.
43. Tsui, I.F.L., Poh, C.F., Garnis, C., Rosin, M.P., Zhang, L. and Lam, W.L. (2009) Multiple pathways in the FGF signaling network are frequently deregulated by gene amplification in oral dysplasias. *Int. J. Cancer*, **125**, 2219–2228.
44. Duran, A., Linares, J.F., Galvez, A.S., Wikenheiser, K., Flores, J.M., Diaz-Meco, M.T. and Moscat, J. (2008) The signaling adaptor p62 is an important NF-kappaB mediator in tumorigenesis. *Cancer Cell*, **13**, 343–354.
45. Shimada, K., Uzawa, K., Kato, M., Endo, Y., Shiiba, M., Bukawa, H., Yokoe, H., Seki, N. and Tanzawa, H. (2005) Aberrant expression of RAB1A in human tongue cancer. *Br. J. Cancer*, **92**, 1915–1921.
46. Bos, P.D., Zhang, X.H.-F., Nadal, C., Shu, W., Gomis, R.R., Nguyen, D.X., Minn, A.J., van de Vijver, M.J., Gerald, W.L., Foekens, J.A. *et al.* (2009) Genes that mediate breast cancer metastasis to the brain. *Nature*, **459**, 1005–1009.
47. Teramoto, T., Kawamata, H. and Fujimori, T. (2003) How does adenosquamous carcinoma arise in the small intestine? *J. Gastroenterol.*, **38**, 810–811.
48. Wang, Z. (2005) The role of COX-2 in oral cancer development, and chemoprevention/treatment of oral cancer by selective COX-2 inhibitors. *Curr. Pharm. Des.*, **11**, 1771–1777.
49. Pontes, H.A.R., Pontes, F.S.C., Fonseca, F.P., de Carvalho, P.L., Pereira, E.M., de Abreu, M.C., de Freitas Silva, B.S. and dos Santos Pinto, D. (2013) Nuclear factor κB and cyclooxygenase-2 immunoexpression in oral dysplasia and oral squamous cell carcinoma. *Ann. Diagn. Pathol.*, **17**, 45–50.
50. Kalyankrishna, S. and Grandis, J.R. (2006) Epidermal growth factor receptor biology in head and neck cancer. *J. Clin. Oncol.*, **24**, 2666–2672.
51. Taoudi Benchekroun, M., Saintigny, P., Thomas, S.M., El-Naggar, A.K., Papadimitrakopoulou, V., Ren, H., Lang, W., Fan, Y.-H., Huang, J., Feng, L. *et al.* (2010) Epidermal growth factor receptor expression and gene copy number in the risk of oral cancer. *Cancer Prev. Res. (Phila)*, **3**, 800–809.
52. Sheu, J.J.-C., Hua, C.-H., Wan, L., Lin, Y.-J., Lai, M.-T., Tseng, H.-C., Jinawath, N., Tsai, M.-H., Chang, N.-W., Lin, C.-F. *et al.* (2009) Functional genomic analysis identified epidermal growth factor receptor activation as the most common genetic event in oral squamous cell carcinoma. *Cancer Res.*, **69**, 2568–2576.
53. Pan, W., Zhang, Q., Xi, Q.S., Gan, R.B. and Li, T.P. (2001) FUP1, a gene associated with hepatocellular carcinoma, stimulates NIH3T3 cell proliferation and tumor formation in nude mice. *Biochem. Biophys. Res. Commun.*, **286**, 1033–1038.
54. Rodríguez, M.I., Peralta-Leal, A., O'Valle, F., Rodriguez-Vargas, J.M., Gonzalez-Flores, A., Majuelos-Melguizo, J., López, L., Serrano, S., de Herreros, A.G., Rodríguez-Manzaneque, J.C. *et al.* (2013) PARP-1 regulates metastatic melanoma through modulation of vimentin-induced malignant transformation. *PLoS Genet.*, **9**, e1003531.
55. Costello, E., Greenhalf, W. and Neoptolemos, J.P. (2012) New biomarkers and targets in pancreatic cancer and their application to treatment. *Nat. Rev. Gastroenterol. Hepatol.*, **9**, 435–444.
56. Onodera, T., Sakai, T., Hsu, J.C., Matsumoto, K., Chiorini, J.A. and Yamada, K.M. (2010) Btbd7 regulates epithelial cell dynamics and branching morphogenesis. *Science*, **329**, 562–565.
57. Wicki, A., Lehembre, F., Wick, N., Hantusch, B., Kerjaschki, D. and Christofori, G. (2006) Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell*, **9**, 261–272.
58. De Vicente, J.C., Rodrigo, J.P., Rodriguez-Santamarta, T., Lequerica-Fernández, P., Allonca, E. and García-Pedrero, J.M. (2013) Podoplanin expression in oral leukoplakia: tumorigenic role. *Oral Oncol.*, **49**, 598–603.
59. Wu, I. and Moses, M.A. (2003) BNF-1, a novel gene encoding a putative extracellular matrix protein, is overexpressed in tumor tissues. *Gene*, **311**, 105–110.
60. Ellermeier, C. and Smith, G.R. (2005) Cohesins are required for meiotic DNA breakage and recombination in Schizosaccharomyces pombe. *Proc. Natl Acad. Sci. USA*, **102**, 10952–10957.
61. Kalejs, M., Ivanov, A., Plakhins, G., Cragg, M.S., Emzinsh, D., Illidge, T.M. and Erenpreisa, J. (2006) Upregulation of meiosis-specific genes in lymphoma cell lines following genotoxic insult and induction of mitotic catastrophe. *BMC Cancer*, **6**, 6.
62. Severino, P., Alvares, A.M., Michaluart, P., Okamoto, O.K., Nunes, F.D., Moreira-Filho, C.A. and Tajara, E.H. (2008) Global gene expression profiling of oral cavity cancers suggests molecular heterogeneity within anatomic subsites. *BMC Res. Notes*, **1**, 113.
63. Nasser, W., Flechtenmacher, C., Holzinger, D., Hofele, C. and Bosch, F.X. (2011) Aberrant expression of p53, p16INK4a and Ki-67 as basic biomarker for malignant progression of oral leukoplakias. *J. Oral Pathol. Med.*, **40**, 629–635.
64. Meric, F. and Hunt, K.K. (2002) Translation initiation in cancer: a novel target for therapy. *Mol. Cancer Ther.*, **1**, 971–979.
65. Weber, A., Wittekind, C. and Tannapfel, A. (2003) Genetic and epigenetic alterations of 9p21 gene products in benign and malignant tumors of the head and neck. *Pathol. Res. Pract.*, **199**, 391–397.
66. Brito, J.A.R., Gomes, C.C., Guimarães, A.L.S., Campos, K. and Gomez, R.S. (2013) Relationship between microRNA expression levels and histopathological features of dysplasia in oral leukoplakia. *J. Oral Pathol. Med.*
67. Yang, C.-C., Hung, P.-S., Wang, P.-W., Liu, C.-J., Chu, T.-H., Cheng, H.-W. and Lin, S.-C. (2011) miR-181 as a putative biomarker for lymph-node metastasis of oral squamous cell carcinoma. *J. Oral Pathol. Med.*, **40**, 397–404.
68. Sadikovic, B., Andrews, J., Carter, D., Robinson, J. and Rodenhiser, D.I. (2008) Genome-wide H3K9 histone acetylation profiles are altered in benzopyrene-treated MCF7 breast cancer cells. *J. Biol. Chem.*, **283**, 4051–4060.
69. Devries, S., Nyante, S., Korkola, J., Segraves, R., Nakao, K., Moore, D., Bae, H., Wilhelm, M., Hwang, S. and Waldman, F. (2005) Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. *J. Mol. Diagn.*, **7**, 65–71.
70. Paris, P.L., Albertson, D.G., Alers, J.C., Andaya, A., Carroll, P., Fridlyand, J., Jain, A.N., Kamkar, S., Kowbel, D., Krijtenburg, P.-J. *et al.* (2003) High-resolution analysis of paraffin-embedded and formalin-fixed prostate tumors using comparative genomic hybridization to genomic microarrays. *Am. J. Pathol.*, **162**, 763–770.
71. Bohlander, S.K., Espinosa, R., Le Beau, M.M., Rowley, J.D. and Díaz, M.O. (1992) A method for the rapid sequence-independent amplification of microdissected chromosomal material. *Genomics*, **13**, 1322–1324.
72. Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*, **25**, 402–408.
73. Brown, K.R. and Jurisica, I. (2005) Online predicted human interaction database. *Bioinformatics*, **21**, 2076–2082.
74. Brown, K.R. and Jurisica, I. (2007) Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol.*, **8**, R95.
75. Shirdel, E.A., Xie, W., Mak, T.W. and Jurisica, I. (2011) NAViGaTing the micronome—using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs. *PLoS ONE*, **6**, e17429.