Next-Gen Sequencing Exposes Frequent MED12 Mutations and Actionable Therapeutic Targets in Phyllodes Tumors

Andi K. Cani1, Daniel H. Hovelson2, Andrew S. McDaniel1, Seth Sadis3, Michaela J. Haller1, Venkata Yadati1, Anmol M. Amin1, Jarred Bratley1, Santoshini Bandla3, Paul D. Williams3, Kate Rhodes4, Chia-Jen Liu1, Michael J. Quist1,5, Daniel R. Rhodes5, Catherine S. Grasso1,5, Venkata Yadati1, Anmol M. Amin1, Jarred Bratley1, Santhoshi Bandla3, Paul D. Williams3, Kate Rhodes4, Chia-Jen Liu1, Michael J. Quist1,5, Daniel R. Rhodes5, Catherine S. Grasso1,5, Celina G. Kleer1, and Scott A. Tomlins1,6,7

Abstract

Phyllodes tumors are rare fibroepithelial tumors with variable clinical behavior accounting for a small subset of all breast neoplasms, yet little is known about the genetic alterations that drive tumor initiation and/or progression. Here, targeted next-generation sequencing (NGS) was used to identify somatic alterations in formalin-fixed paraffin-embedded (FFPE) patient specimens from malignant, borderline, and benign cases. NGS revealed mutations in mediator complex subunit 12 (MED12) affecting the G44 hotspot residue in the majority (67%) of cases spanning all three histologic grades. In addition, loss-of-function mutations in p53 (TP53) as well as deleterious mutations in the tumor suppressors retinoblastoma (RB1) and neurofibromin 1 (NF1) were identified exclusively in malignant tumors. High-level copy-number alterations (CNA) were nearly exclusively confined to malignant tumors, including potentially clinically actionable gene amplifications in IGF1R and EGFR. Taken together, this study defines the genomic landscape underlying phyllodes tumor development, suggests potential molecular correlates to histologic grade, expands the spectrum of human tumors with frequent recurrent MED12 mutations, and identifies IGF1R and EGFR as potential therapeutic targets in malignant cases.

Implications: Integrated genomic sequencing and mutational profiling provides insight into the molecular origin of phyllodes tumors and indicates potential druggable targets in malignant disease.

Visual Overview: http://mcr.aacrjournals.org/content/early/2015/04/02/1541-7786.MCR-14-0578/F1.large.jpg


Introduction

Phyllodes tumors of the breast are relatively rare fibroepithelial tumors that account for approximately 1% of all breast neoplasms. Like benign breast fibroadenomas, they are characterized by proliferation of both stromal and epithelial components, but, in contrast, they have considerable malignant potential. Phyllodes tumors are classified as benign (~65%), borderline (~25%), and malignant (~10%) based on histologic features, including cellular atypia, mitotic activity, stromal overgrowth, stromal cellularity, and tumor margins (1). However, this histopathologic classification often fails to predict which phyllodes tumors will recur or metastasize after treatment and does not accurately inform on treatment options. Although local recurrence after resection is most prevalent in histologically malignant cases (approximately 30%, depending on width of excised margins), borderline, and benign tumors can also recur locally in about 15% and 10% of cases, respectively, demonstrating the limitations of current prognostic approaches (2). Likewise, although approximately 10% of all phyllodes tumors progress to distant metastases, only approximately 20% of histologically malignant cases do so (3, 4), leaving a substantial number of borderline and even histologically benign cases that have metastatic potential. Conversely, although most histologically benign cases will behave as such, there are a proportion of phyllodes tumors classified as malignant and border-line that will behave in a benign manner. Current treatment guidelines for phyllodes tumors require wide surgical resection margins, but efficacious treatment options for the 10% of all phyllodes tumors that progress to metastatic disease are lacking and survival rates are dismal (3).

The key genetic alterations driving phyllodes tumor development and molecular correlates with histologic grade and malignant behavior are poorly characterized. Comparative genomic hybridization (CGH) and array CGH (aCGH) studies...
have shown multiple recurrent, broad somatic chromosomal copy-number alterations (CNA) in phyllodes tumors, including gains of chromosome 1q and losses in 13q, 6q, 9p; however, their prognostic utility is unclear (5–9). Several genes have been implicated in phyllodes tumor development by virtue of being localized to areas of CNA, including *EGFR*, which was recently shown by FISH to be amplified in 2% to 16% of cases (10, 11). In addition, gene expression and IHC studies have implicated various signaling pathways, including insulin-like growth factor (IGF) and Wnt/β-catenin, as being activated in phyllodes tumors (1). To more comprehensively assess somatic molecular alterations in phyllodes tumors and identify potential opportunities for personalized medicine, we performed next-generation sequencing (NGS) of 15 formalin-fixed paraffin-embedded (FFPE) phyllodes tumors representing the histologic grade spectrum.

### Materials and Methods

#### Case selection

We identified a cohort of 15 archived, routine clinical FFPE phyllodes tumor specimens from the University of Michigan, Department of Pathology Tissue Archive. Clinicopathologic information for each case was obtained from the clinical archive. Hematoxylin and eosin (H&E)–stained slides for all cases were reviewed by a board-certified Anatomic Pathologist (S.A. Tomlins) to ensure sufficient tumor content and confirm histologic grade.

#### Targeted next-generation sequencing

Targeted NGS of tumor tissue was performed with IRB approval. For each specimen, 4 to 10 × 10^6 FFPE sections were cut from a single representative block per case, using macrodissection with a scalpel as needed to enrich for at least 50% tumor content (as defined by areas of stromal overgrowth). DNA was isolated using the Qiagen Allprep FFPE DNA/RNA Kit (Qiagen), according to the manufacturer’s instructions except for adding a 2 minute room temperature incubation and extending centrifugation time to 5 minutes during the xylene deparaffinization (step 1) and ethanol washing of xylene (step 2). DNA was quantified using the Qubit 2.0 fluorometer (Life Technologies).

Targeted, multiplexed PCR-based NGS was performed on each component using a custom panel comprised of 2,462 amplicons targeting 130 genes and Ion Torrent–based sequencing. Genes included in this panel were selected based on pan-cancer NGS and copy-number profiling data analysis that prioritized somatic, recurrently altered oncogenes, tumor suppressors, genes present in high-level copy gains/losses, and known/investigational therapeutic targets. Barcoded libraries were generated from 20 ng of DNA per sample using the Ion Ampliseq library Kit 2.0 (Life Technologies) according to the manufacturer’s instructions with barcode incorporation. Templates were prepared using the Ion PGM Template OT2 200 Kit (Life Technologies) on the Ion One Touch 2 according to the manufacturer’s instructions. Sequencing of multiplexed templates was performed using the Ion Torrent Personal Genome Machine (PGM) on Ion 318 chips using the Ion PGM Sequencing 200 Kit v2 (Life Technologies) according to the manufacturer’s instructions.

Data analysis was performed essentially as described previously (12) in Torrent Suite 4.0.2, with alignment by TMAP using default parameters, and variant calling using the Torrent Variant Caller plugin (version 4.0-r76860) using default low-stringency somatic variant settings. Variants were annotated using Annovar (13). Called variants were filtered to remove synonymous or noncoding variants, those with flow corrected read depths (FDP) less than 20, flow corrected variant allele containing reads (FAO) less than 6, variant allele frequencies (FAO/FDP) less than 0.10, extreme skewing of forward/reverse flow corrected reads calling the variant (FAO/FSAR <0.2 or >5, or FAO or FSAR <3), or indels within homopolymer runs >4. Variants occurring exclusively in reads containing other variants (single-nucleotide variants or indels) or those occurring in the last mapped base of a read were excluded. Variants with allele frequencies >0.5% in ESP6500 or 1,000 genomes or those reported in ESP6500 or 1,000 genomes with observed variant allele frequencies between 0.40 and 0.60 or >0.9 were considered germ line variants. High-confidence somatic variants passing the above criteria were then visualized in IGV. We have previously confirmed that these filtering criteria identify variants that pass PCR validation with >95% accuracy (12). To prioritize potential driving alterations, we used Oncomine software tools (powertools.oncomine.com) to annotate called variants, which uses pan-cancer NGS data to identify genes as oncogenes or tumor suppressors, based on overrepresentation of hotspot or deleterious mutations, respectively. Variants in oncogenes are then considered gain-of-function if at a hotspot and variants in tumor suppressors are considered loss-of-function if deleterious or at a hotspot (S.A. Tomlins; unpublished data).

#### Copy-number analysis

To identify CNAs, normalized, GC-content corrected read counts per amplicon for each sample were divided by those from a pool of normal male genomic DNA samples (FFPE and frozen tissue, individual, and pooled samples), yielding a copy-number ratio for each amplicon. Gene-level copy-number estimates were determined by taking the coverage-weighted mean of the per-amplon ratios, with expected error determined by the probe-to-probe variance (12); a detailed article describing this technique is in submission (C.S. Grasso; submitted for publication). Genes with a log2 copy-number estimate of <-1 or >0.6 were considered to have high-level loss or gain, respectively.

#### Sanger sequencing to validate called somatic variants

Bidirectional Sanger sequencing was performed over the observed *MED12* mutation hotspot (G44) on all tumor samples. Genomic DNA (10 ng) was used as template in PCR amplifications with Invitrogen Platinum PCR Supermix Hi-Fi (Life Technologies) with the suggested initial denaturation and cycling conditions. Primer sequences were as previously reported (14, 15) with the addition of universal M13 adaptors (M13 forward: TGTTAAAACGACGGCCAGT and M13 reverse: CAG-GAAAACGAGCTATGACC). PCR products were subjected to bidirectional Sanger sequencing for both primer pairs by the University of Michigan DNA Sequencing Core after treatment with ExoSAP-IT (GE Healthcare) and sequences were analyzed using SeqMan Pro software (DNASTAR).

#### qPCR to validate copy-number variations

*EGFR*, *IGF1R*, and *CDKN2A* copy-number changes were subjected to validation through quantitative real-time PCR (qPCR) for 12 samples with sufficient DNA. PH13, 14 and 30 had insufficient DNA (no CNAs in these genes were identified by NGS) for qPCR and PH5 had sufficient DNA only for assessing...
Next-Generation Sequencing of Phyllodes Tumors

**EGFR** and **CDKN2A** (no **IGF1R** CNAs were identified by NGS). Primers and probes (5’ FAM; ZEN/Iowa Black FQ dual quenchers) were designed using PrimerQuest (http://www.idtdna.com/Primerquest/Home/Index, hg 19 genome assembly) and obtained from IDT. Assay specificity was confirmed using BLAST and BLAT and primers/probes in areas of SNPs were excluded. Primer/probe sequences are available upon request.

BLAST and BLAT and primers/probes in areas of SNPs were excluded. Primer/probe sequences are available upon request.

**Results**

We performed targeted NGS on a cohort of 15 FFPE phyllodes tumors comprised of 5 cases each of benign, borderline, and malignant histologic grade; representative photomicrographs and clinical characteristics of all patients are presented in Fig. 1A.

**Figure 1.**
Histology and clinicopathologic information for FFPE phyllodes tumors assessed by targeted NGS and integrative molecular heatmap of driving molecular alterations in phyllodes tumors. A, H&E-stained sections of representative benign (PH-19), borderline (PH-08), and malignant (PH-05 and PH-06) phyllodes tumors sequenced are shown. Top, ×4 original magnification; bottom, ×20 magnification. Clinicopathologic information, including histologic grade, specimen type, patient age, procedure type, and tumor size, for all cases is given (Ex./Lump, excisional biopsy or lumpectomy; Mast, mastectomy; Core bx, Core biopsy). B, targeted NGS of 15 FFPE phyllodes tumors was performed to identify potentially driving/actionable molecular alterations. All high-confidence, gain- or loss-of-function somatic mutations in oncogenes and tumor suppressors, in addition to high-level CNA are indicated for each case. Specific alteration types are indicated according to the legend (Nonsense SNV, Nonsynonymous SNV, Fs and Fp indel, frame-shifting and frame-preserving indels, respectively). Slashed boxes indicate two alterations. Clinicopathologic information is shown above the heatmap according to the legend and as in A.
We isolated an average of 0.65 μg DNA per case from 4 to 10 × 10-μm sections using microdissection to enrich tumor content as needed. NGS was performed using a multiplexed PCR-based custom Ion Torrent Ampliseq panel comprising of 2,462 amplicons targeting 130 genes and Ion Torrent–based sequencing on the PGM. Targeted genes were selected on the basis of pan-cancer NGS and copy-number profiling data analysis to prioritize somatic, recurrently altered oncogenes, tumors suppressors and genes present in high-level CNAs. Detailed characterization of this panel will be reported separately (S.A. Tomlins; unpublished data).

NGS of multiplexed templates on the Ion Torrent PGM generated an average of 1,011,571 mapped reads yielding >409 targeted base coverage across the 15 samples (Supplementary Table S1). We identified a total of 26 high-confidence, likely somatic nonsynonymous or splice site altering point mutations and short insertion/deletions (indels) across the 15 samples (median 2; range, 0–4) as shown in Supplementary Tables S2 and S3. The number of high-confidence somatic nonsynonymous mutations were not significantly different between the histologic grades (Kruskal–Wallis test, \( P = 0.09 \)), as shown in Supplementary Fig. S1A. Copy-number analysis of NGS data yielded a total of 16 high-level CNAs (median 0; range, 0–6). The number of high level CNAs differed significantly between histologic grades (Kruskal-Wallis test, \( P = 0.002 \)), with malignant tumors harboring significantly more high-level CNAs per sample (median 2; range, 0–6) than borderline (median 0; range, 0–0) or benign (median 0; range, 0–2) tumors (Kruskal–Wallis test, post-hoc analysis, both \( P < 0.05 \)), as shown in Supplementary Fig. S1B. Prioritized likely gain- or loss-of-function somatic mutations in oncgenes and tumor suppressors (see below) and high-level CNAs for each case are shown in an integrative heatmap (Fig. 1B).

By NGS, we found that MED12, which encodes subunit 12 of the Mediator complex (the multiprotein assembly that serves as a general coactivator of transcription by RNA polymerase II) was mutated in 10 of 15 samples (67%; one sample with biallelic mutations) by automated variant calling and visual read inspection in IGV (as some called variants were filtered due to skewed read support). All mutations were localized to the exon 2 hotspot region near residue G44 (Fig. 2A and Supplementary Tables S2 and S3), which has recently been reported to be recurrently mutated at high frequency in uterine leiomyomas (14–16) and benign breast fibroadenomas (17), and more rarely in uterine leiomyosarcomas (15, 16, 18, 19). Five of 11 total MED12 mutations were point mutations at G44 (3 p.G44S, 1 p.G44C, and 1 p.G44R) whereas three mutations were frame-preserving deletions adjacent to or including G44 (p.38_43, p.41_49, and p.42_51). Two mutations were intronic point mutations just upstream of exon 2, at a previously reported splice acceptor site causing retention of an additional 6 bases in the transcript (c.IVS-8 p.E33_D34insPQ; refs. 14, 17). PH-11, which harbored a c.IVS-8 mutation, also harbored an additional intronic mutation further upstream (c.IVS-15), consistent with biallelic intronic MED12 mutations in this sample. There was no significant difference in the presence of MED12 mutations between tumors of different histologic grade (benign 4/5, borderline 4/5, malignant 2/5, Fisher exact test, \( P = 0.5 \)). All MED12 mutations were confirmed by bidirectional Sanger sequencing (Fig. 2B).

To prioritize potential driving alterations from the remaining non-MED12 point mutations/indels, we used the Oncome Plugin in Ion Reporter for assessing gain- or loss-of-function. This analysis identified five loss-of-function alterations, including three mutations in TP53 (T270L in PH-14, Q192X in PH-16, and C242Y in PH13), and one mutation each in RBL (E533X in PH16) and NF1 (p.1152_1153del in PH-05), as shown in the integrative heatmap of driving alterations (Fig. 1B and Supplementary Tables S2 and S3). Intriguingly, these loss-of-function alterations occurred exclusively in malignant tumors.

Copy-number analysis of NGS data demonstrated recurrent low-level CNAs, including gain of chromosome 1q and loss of chromosome 13q, consistent with previous reports (5–9). These were more prevalent in malignant tumors (5/5), but were also present in two borderline and one benign case (Fig. 3A and Supplementary Fig. S2). High-level CNAs were nearly exclusively present in malignant specimens 14 of 16 alterations, as shown in Fig. 3A and B. Of note, PH-03 showed high-level EGFR (copy-number ratio > 6) and IGF1R amplifications, whereas PH-16 also showed a high-level IGF1R amplification (copy-number ratio > 32, Fig. 3B). TERT amplifications were also observed in three malignant tumors, whereas PH-05 harbored a focal high-level CDKN2A (p16INK4A) loss. We confirmed EGFR, IGF1R, and CDKN2A CNAs by qPCR as shown in Fig. 3C.

Discussion

We performed targeted NGS of 15 FFPE phyllodes tumors representing all three histologic grades to identify somatic alterations associated with tumor development and potential targetable alterations. Mutations in MED12 were present in 10 of 15 cases (67%) and affected the known exons 2 G44 residue hotspot through multiple mechanisms. No significant difference in MED12 mutation frequency was observed across histologic grades, although our cohort size is limited. Our IRB approved protocol does not allow NGS of matched normal tissue; however, the observed MED12 variant allele frequencies are consistent with somatic events as seen in other tumors. Similar MED12 somatic mutations are frequent (50%–70%) in benign uterine leiomyomas (14–16), but less common in malignant uterine leiomyosarcomas (7%–30%; refs. 15, 16, 18, 19). Recently, Lim and colleagues identified similar MED12 mutations in 59% of benign breast fibroadenomas through exome sequencing (17). Given the morphologic similarity of breast fibroadenomas and benign phyllodes tumors, frequent MED12 mutations in both entities support a closely related molecular origin. In addition, although our findings will need to be replicated in larger cohorts, the similar frequency of MED12 mutations across the histologic spectrum of phyllodes tumors (in addition to benign fibroadenomas) suggests that MED12 mutations in the breast are early events that may be unrelated to malignant behavior, in contrast with uterine leiomyomas and leiomyosarcomas, which show notable differences in MED12 mutation frequencies as just described. Our results also support the evolution of malignant phyllodes tumors from less aggressive fibroadenomas or phyllodes tumors (possibly through loss of key tumor suppressors). Although MED12 hotspot mutations have been identified infrequently in extraterine or extramammary tumors (18, 20), functional studies support a role for MED12 mutations affecting the G44 hotspot in dysregulation of estrogen signaling in estrogen responsive cells (17), and the Mediator complex is known to interact with the estrogen receptor. (21).
Although surgical resection of phyllodes tumors may be curative, local recurrence is not uncommon and distant metastasis is associated with poor survival. Furthermore, the histologic features do not accurately predict clinical behavior of phyllodes tumors. Hence, targetable alterations, particularly in malignant phyllodes tumors, may be useful for personalized medicine strategies. Through copy-number analysis of NGS data (and confirmed by qPCR), we identified potentially clinically actionable high-level, focal amplifications of EGFR and IGF1R in 7% and 13% of cases, respectively (1/5 and 2/5 malignant cases). EGFR has been shown to be highly amplified in phyllodes tumors by FISH in up to 16% of cases (10, 11), consistent with our findings. Dysregulation of the IGF pathway has been implicated in phyllodes tumors by IHC (1); however, IGF1R amplification has not been reported.

Direct comparison of additional CNAs identified in our study and previous studies is challenging due to platform differences. However, broad, low-level gains in genes on 1q and losses on 13q were observed in malignant as well as borderline and benign tumors, consistent with previous...
Cani et al.

Figure 3. Copy-number analysis of phyllodes tumor identifies potential therapeutic targets in malignant samples. Copy-number analysis was performed from NGS data. For each sequenced phyllodes tumor, GC content corrected, normalized read counts per amplicon were divided by those from a composite normal sample, yielding a copy-number ratio for each amplicon. Gene-level copy-number estimates were determined by taking the weighted mean of the per-probe copy-number ratios. A, summary of gene-level copy-number ratios (log2) for all profiled samples. Selected genes of interest with high-level CNAs are colored according to the legend. B, copy-number profiles for three malignant phyllodes tumors with high-level CNAs. Log2 copy-number ratios per amplicon are plotted (with each individual amplicon represented by a single dot, and each individual gene indicated by different colors), with gene-level copy-number estimates (black bars) determined by taking the weighted mean of the per-probe copy-number ratios. Selected high-level CNAs are indicated. C, dPCR confirmation of high-level CNAs in EGFR, IGF1R, and CDKN2A. dPCR on genomic DNA from indicated samples was performed using DNMT3A, FBXW7, and MYO18A as the reference genes. Normalized mean IGF1R (blue), EGFR (red), and CDKN2A (green) log2 copy-number ratios [using PH22 (no CNAs by NGS) as the calibrator] from triplicate qPCR ± SD are plotted.

reports. On the other hand, we did not observe gains in MDM2 or MDM4, which have been reported in previous aCGH studies, and were targeted herein; we hypothesize this may be due to the high TP53 alteration rate in our phyllodes tumors with high numbers of CNAs. Finally, our panel did not target genes in some previously reported regions of CNA (such as 6q), precluding comparisons of these alterations.

Besides MED12 hotspot mutations, other potential driving somatic point mutations/indels, which included loss-of-function alterations in TP53, RB1, and NF1, occurred exclusively in malignant tumors. In addition, high-level, focal CNAs (such as those in EGFR and IGF1R) were only observed in malignant cases. Together, these findings support molecular correlates to histologic grade. Whether such molecular alterations may be useful in cases with challenging histology or show prognostic potential can be investigated in additional cohorts.

Taken together, our results demonstrate frequent MED12 mutations in phyllodes tumors, supporting a shared origin with benign breast fibroadenomas. In addition, our results suggest potential therapeutic targets in malignant tumors, including EGFR and IGF1R. Finally, as both driving somatic mutations/indels other than MED12 and high-level, focal CNAs occurred exclusively in malignant tumors in our cohort, such alterations may be useful for classification or prognostication in borderline tumors if confirmed in other cohorts.

Disclosure of Potential Conflicts of Interest
S.A. Tomlins has a separate sponsored research agreement with Compendia Bioscience/Life Technologies/ThermoFisher Scientific that provides access to the sequencing panel used herein. No other aspect of the study described herein was supported by Compendia Bioscience/Life Technologies/ThermoFisher Scientific. S. Sadis, S. Bandla, P.D. Williams, K. Rhodes,
and D.R. Rhodes are employees of ThermoFisher Scientific. The other authors have no competing interests to declare.

**Authors’ Contributions**

Conception and design: A.K. Cani, A.S. McDaniel, K. Rhodes, D.R. Rhodes, S.A. Tomlins
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.K. Cani, A.S. McDaniel, V. Yadati, A.M. Amin, J. Bratley, K. Rhodes
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.K. Cani, D.H. Hovelson, M.J. Quist, C.S. Grasso, C.G. Kleer, S.A. Tomlins

Writing, review, and/or revision of the manuscript: A.K. Cani, D.H. Hovelson, A.S. McDaniel, S. Sadis, M.J. Haller, A.M. Amin, C.S. Grasso, C.G. Kleer, S.A. Tomlins

**References**


