Mutational profiling of lipomas

Deepika Kanojia^{1*}, Pushkar Dakle¹, Anand Mayakonda^{1,2}, Rajeev Parameswaran³, Mark E Puhaindran⁴, Victor Lee Kwan Min⁵, Vikas Madan¹ and H Phillip Koeffler^{1,6,7}

¹Cancer Science Institute of Singapore, National University of Singapore, Singapore; ²Epigenomics and Cancer Risk Factors, German Cancer Research Center (DKFZ), Heidelberg, Germany; ³Division of Surgical Oncology, National University Cancer Institute, Singapore; ⁴Department of Hand and Reconstructive Microsurgery, National University Hospital; ⁵Department of Pathology, National University Hospital; ⁶Division of Hematology/Oncology, Cedars-Sinai Medical Center, University of California, School of Medicine, Los Angeles, California, USA; ⁷National University Cancer Institute, National University Hospital, Singapore.

***Correspondence:** csidk@nus.edu.sg

Abstract

Lipomas are benign fatty tumors with a high prevalence rate, mostly found in adults but have a good prognosis. Until now, reason for lipoma occurrence not been identified. We performed whole exome sequencing to define the mutational spectrum in ten lipoma patients along with their matching control samples. We identified 412 somatic variants including missense mutations, splice site variants, frameshift indels, and stop gain/lost. Kinase genes and transcriptions factors were among the validated mutated genes critical for cell proliferation and survival. Pathway analysis revealed enrichment of calcium, Wnt and phospholipase D signaling in patients. Whole exome sequencing in lipomas identified mutations in genes with a possible role in development and progression of lipomas.

Keywords

Lipoma, whole exome sequencing, mutation signature

Author Summary

We presented genomic insight into the development of lipomas, the most common benign tumor of soft tissue. Until date, no one knows the cause of lipoma development and its progression. Our group for the first time profiled ten lipoma patients' samples and their matching normal controls to delineate the somatic mutation pattern using whole exome sequencing. Interestingly, exome-profiling results highlighted the potential role of important kinase genes and transcription factors in lipoma development. In addition, calcium, Wnt and phospholipase D signaling pathway speculated to be involve in pathogenesis of this disease.

Introduction

Lipomas, benign soft tissue fatty tumors, are the most common benign mesenchymal tumors [1]. Lipomas are slow growing tumors and often occur under the skin on the neck, shoulders, arms, back, abdomen and thighs. However, occasionally lipomas may be present in deep location or originate within muscle [2]. Lipoma affects only 1% of the population, although it is probably underreported [3]. These benign tumors are commonly found in adults and their prevalence is higher in men than in women. Lipomas are generally small round masses of less than 5 cm and typically are harmless unless they compress an organ. Patients may have either a solitary lipoma or multiple lipomas. Lipomas commonly are diagnosed by clinicians and histologically examined after surgical resection. Interestingly, prognosis for lipomas is very good with almost no chance of recurrence after removal of the tumor [4].

Exact cause of lipomas is still not known, but lipomas have been identified with genetic rearrangements including structural changes at 12q13-15, 13q and 6p21-23 regions [5]. Approximately, 55%-75% of solitary lipomas have cytogenetic abnormalities involving *HMGA2* gene rearrangements (truncation or fusion) [6]. Studies revealed enhanced adipogenesis and a normal rate of adipocyte apoptosis in lipoma tissue compared to normal fat tissue [7].

A more aggressive form of fatty tumor is liposarcoma, which is rare in occurrence but they frequently develop dangerous metastatic tumors which is symptomatically different from lipomas [4]. Whole exome sequencing analysis on different subtypes of liposarcoma identified driver somatic mutations including *PI3KCA*, *TP53*, *CHEK2*, *NF1*, *ATM*, *BRCA1* and *EGFR* [8]. Genomic characterization of tumors improved our understanding of their molecular genesis for better management of the disease.

We screened superficial subcutaneous lipomas using whole exome sequencing to systematically profile their mutational signatures, and identify affected pathways to understand further their molecular pathogenesis.

Results and discussion

We performed a genomic analysis to profile the somatic mutational landscape of lipomas by whole exome sequencing. To best of our knowledge, this is first study to perform exome sequencing on lipomas.

Clinical characteristics of lipoma patients

Primary subcutaneous lipomas were obtained from patients with their consent undergoing surgical resection at the National University Hospital. Detailed clinical features associated with these benign tumors presented in Supplementary Table 1. We collected lipomas from ten patients (all males) with a median age of 52 years. We performed whole exome sequencing on ten lipomas and normal controls from the same patient (either blood sample or adjacent normal tissue).

Sequencing data

Whole exome sequencing of lipoma samples achieved an average of 135 million reads per sample (range 111-188 million). The control samples from the same patients found to have an average of 128 million reads per sample (range 112-164 million). Sequencing obtained a mean depth of 93 and 85 across the sequencing regions in lipomas and matching controls, respectively.

Variant allele frequency (VAF) implicate how early a mutation occurred during progression of the tumor and our data suggest that the mutations occurred both early (variants

with high VAF) and later (variants with low VAF) during development of their lipoma. In total, 5 frameshift deletions and 5 frameshift insertions occurred with 5-20% VAF.

Mutational landscape in lipomas

Using our data pipeline analysis, we identified 412 somatic variants with a median of 42 per patient sample (range 11-79) [Fig 1A and Supplementary Tables 2-11]. The type of mutations include 367 missense mutations, 19 stop gained, 5 stop lost, 9 splice acceptor, 2 splice site donor, 5 frameshift insertions and 5 frameshift deletions [Fig 1B]. Details of types of mutation for each sample listed in Supplementary Table 12. Similar to other benign tumors and sarcomas, overall mutation burden in lipomas observed to be relatively low. A total of 42 variants were confirmed to be mutated by Sanger sequencing including 37 missense mutations, 1 frameshift deletion, 2 in-frame deletions, 1 splice donor and 1 stop gained mutation. Detailed list of variants, their mutation type and known function of genes provided in Supplementary Tables 13. Validated mutations included kinase genes; ACVRIC (VAF 28%), PTPRT (VAF 14%), ERBB2 (VAF 14%) and MAPK7 (VAF 20%) and transcription factors ATMIN (VAF 7%), ZNF317 (VAF 26%) and FOSL1 (VAF 19%); each are crucial for cell survival and proliferation. No single mutated gene was common in all ten tumors. Several of the genes are known to play functional role in oncogenesis such as APC (VAF 26%), FOSL1 (VAF 19%), ERBB2 (VAF 14%), NFATC3 (VAF 10%), PDE1A (VAF 11%), ATMIN (VAF 7%), and MAPK7 (VAF 20%). In addition, experimental evidence have also indicated modulation of FOSL1 (VAF 19%) [9], ERBB2 (VAF 14%) [10], and PDE1A (VAF 11%) [11] regulates adipogenesis.

Mutational patterns in lipoma patient samples

We performed mutational spectrum analysis of lipomas to categorize their mutational signature and to identify functional mutagenic processes in lipomas. Analysis revealed mutational

signature 15 as the major signature affecting lipoma patients [Figs 2A and 2B]. Signature 15 is one of the mutational signature associated with defective DNA mismatch repair. Patient wise signature analysis is shown in Supplementary Fig1.

Altered pathways in lipomas

Pathway analysis was conducted to identify pathways and biological processes affected by mutated genes in lipomas. This analysis included gene variants validated and confirmed with Sanger sequencing. KEGG analysis indicated that out of 42 validated genes, eleven were involved in signalling of calcium ($P = 0.0005$), Wnt ($P = 0.003$), phospholipase D ($P = 0.03$) as well as stem cells' pluripotency ($P = 0.003$) and axon guidance ($P = 0.03$) pathway [Supplementary Fig2A]. In addition, biological processes were identified which were associated with genes mutated in lipoma samples including calcium ion transport and microtubule regulation as well as signaling through the insulin receptor and tyrosine kinase [Supplementary Fig2B]. Interestingly, calcium signalling pathway and calcium ion transport biological process emerged as a major pathway/process affected by mutated genes in lipoma. Several studies have indicated a regulatory role of calcium ions in adipogenesis contributing to increased fat tissue by promoting and accelerating pre-adipocyte differentiation [12, 13], which is a feature of lipomas. In addition, Wnt signaling [14] and phospholipase D signalling [15] pathways shown to regulate adipogenesis.

Conclusions

We identified and validated mutations in ten lipoma patient samples using whole exome sequencing analysis. However, recurrent gene variants were not noted in ten lipoma patients; thus these alterations could be passenger mutations. Additional samples need to be sequenced and the functional consequences of the identified mutations need to be explored to support the present findings. Present analysis also delineated lipoma-specific mutational signatures and

identified wide-ranging genomic landscape of altered genes and pathways in lipomas that could be useful to investigate further the functional cause of lipoma and its progression.

Materials and Methods

Tissue collection and sample preparation

Lipomas, normal adjacent control tissues and patient's blood samples were obtained from National University Hospital, Tissue repository, Singapore in accordance with ethical guidelines of NUH Institutional Review Board. According to their requirement, samples were collected after obtaining informed written consent from patients. Tissue samples or blood received after surgery was snap frozen and stored in liquid nitrogen until processing. Genomic DNAs were isolated from tumors, control tissues and blood samples using QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's instructions. Genomic DNAs were quantified by Qubit dsDNA BR assay kit (Life technologies) and their quality verified by agarose gel electrophoresis.

Whole exome sequencing

Exome libraries were prepared using the Agilent SureSelect All Human Exon v5 capture kit [8]. Briefly, the qualified genomic DNA samples were randomly fragmented into base pair peak of 150 to 200bp, and adapters ligated to both ends of the resulting fragments. Adapter-ligated templates were purified by the Agencourt AMPure SPRI beads, and fragments with insert size of about 200 bp were excised. Extracted DNA was amplified by ligation-mediated polymerase chain reaction (LM-PCR), purified, and hybridized to the SureSelect Biotinylated RNA Library (BAITS) for enrichment. Hybridized fragments were bound to the streptavidin beads whereas non-hybridized fragments were washed away after 24h. Captured LM-PCR

products were subjected to Agilent 2100 Bioanalyzer to estimate the magnitude of enrichment. Each captured library was independently loaded on Hiseq4000 platform for high-throughput sequencing to ensure that each sample meets the desired average fold-coverage. Raw image files were processed by Illumina base calling Software 1.7 with default parameters, and the sequences were generated as 90bp paired-end reads.

Analysis of whole exome sequencing

Paired end sequencing reads were aligned to the human reference genome (hg19) using bwa mem. Samblaster was subsequently used to mark PCR duplicates [16]. Systematic errors in base quality scores were then corrected using GATK4 Base Recalibrator [17]. Somatic single nucleotide variants and short indels were detected against a matched normal using a consensus calling approach with both GATK4 Mutect2 and Strelka2 [18]. Variants calls from Mutect2 further filtered with GATK4 FilterMutectCalls. Vcf2maf was used to convert the .vcf files to MAF format and also annotated using Ensembl VEP (Version 92_GRCh37) [19]. Depth of coverage for all samples calculated using GATK3 DepthOfCoverage.

We excluded any variants which were not tagged PASS by FilterMutectCalls3 and Strelka2. Variants were further filtered to remove those with global minor allele frequency (GMAF) lower than 1% or variants tagged as common variants by vcf2maf based on allele count in Exac populations, while still retaining those classified as “pathogenic” in ClinVar. Variants with either LOW/MODIFIER impact as per VEP or with variant allele frequency (VAF) less than 0.01 were also removed.

Each mutation categorized into seven categories AT transitions, AT transversions, CpG transitions, CpG transversions, CG transitions/ transversions and Indels. MuSiCa, online tool was used for somatic characterization of mutation signatures present in the patient’s variants list [20]. Refseq-annotated mutations (genes) were used for pathway analysis using PathScan algorithm integrated into MuSic pipeline [21, 22].

Sanger sequencing validations

Gene variants with VAF greater than 10% were validated by Sanger sequencing. Polymerase chain reaction was performed using 25 ng template DNA of lipoma tissue and normal control sample from the same individual. PCR products run on 1% agarose gel and purified using Wizard® SV Gel and PCR Clean-Up System (Promega). Purified PCR products were sequenced and analyzed with the Sequence Scanner Software 2 Version 2 (Applied Biosystems).

Declarations

Funding

This study is supported by the National Institute of Health (R01-CA200992-03), National Research Foundation Singapore under its Singapore Translational Research (STaR) Investigator Award (NMRC/STaR/0021/2014) and administered by the Singapore Ministry of Health's National Medical Research Council (NMRC), the NMRC Centre Grant awarded to National University Cancer Institute of Singapore, the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiatives and is additionally supported by the Wendy Walk Foundation.

Availability of data and materials

All data generated during this study are included in the published article and its supplementary tables.

Authors' contributions

DK and HPK designed the experiments; DK performed the experiments; DK, DPP, AMT and VM analyzed the results; RP, MEP and VLK provided the patient samples; HPK supervised the project and DK, DDP and HPK wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Patient's tissues and blood samples collected in accordance with ethical guidelines of NUH Institutional Review Board and after obtaining informed written consent from patients.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Kransdorf MJ, Moser RP, Jr., Meis JM, Meyer CA. Fat-containing soft-tissue masses of the extremities. *Radiographics*. 1991;11(1):81-106. Epub 1991/01/11. doi: 10.1148/radiographics.11.1.1996399. PubMed PMID: 1996399.
2. McTighe S, Chernev I. Intramuscular lipoma: a review of the literature. *Orthop Rev (Pavia)*. 2014;6(4):5618. Epub 2015/01/09. doi: 10.4081/or.2014.5618. PubMed PMID: 25568733; PubMed Central PMCID: PMC4274454.
3. Luba MC, Bangs SA, Mohler AM, Stulberg DL. Common benign skin tumors. *Am Fam Physician*. 2003;67(4):729-38. Epub 2003/03/05. PubMed PMID: 12613727.
4. Dalal KM, Antonescu CR, Singer S. Diagnosis and management of lipomatous tumors. *J Surg Oncol*. 2008;97(4):298-313. Epub 2008/02/21. doi: 10.1002/jso.20975. PubMed PMID: 18286473.
5. Nishio J. Contributions of cytogenetics and molecular cytogenetics to the diagnosis of adipocytic tumors. *J Biomed Biotechnol*. 2011;2011:524067. Epub 2011/01/29. doi: 10.1155/2011/524067. PubMed PMID: 21274402; PubMed Central PMCID: PMC3025394.
6. Hatano H, Morita T, Ogose A, Hotta T, Kobayashi H, Segawa H, et al. Clinicopathological features of lipomas with gene fusions involving HMGA2. *Anticancer Res*. 2008;28(1B):535-8. Epub 2008/04/04. PubMed PMID: 18383898.
7. Suga H, Eto H, Inoue K, Aoi N, Kato H, Araki J, et al. Cellular and molecular features of lipoma tissue: comparison with normal adipose tissue. *Br J Dermatol*. 2009;161(4):819-25. Epub 2009/06/30. doi: 10.1111/j.1365-2133.2009.09272.x. PubMed PMID: 19558598.
8. Kanojia D, Nagata Y, Garg M, Lee DH, Sato A, Yoshida K, et al. Genomic landscape of liposarcoma. *Oncotarget*. 2015;6(40):42429-44. Epub 2015/12/09. doi:

10.18632/oncotarget.6464. PubMed PMID: 26643872; PubMed Central PMCID: PMCPMC4767443.

9. Luther J, Driessler F, Megges M, Hess A, Herbort B, Mandic V, et al. Elevated Fra-1 expression causes severe lipodystrophy. *J Cell Sci.* 2011;124(Pt 9):1465-76. Epub 2011/04/14. doi: 10.1242/jcs.079855. PubMed PMID: 21486951.

10. Harrington M, Pond-Tor S, Boney CM. Role of epidermal growth factor and ErbB2 receptors in 3T3-L1 adipogenesis. *Obesity (Silver Spring).* 2007;15(3):563-71. Epub 2007/03/21. doi: 10.1038/oby.2007.562. PubMed PMID: 17372305.

11. Snyder PB, Esselstyn JM, Loughney K, Wolda SL, Florio VA. The role of cyclic nucleotide phosphodiesterases in the regulation of adipocyte lipolysis. *J Lipid Res.* 2005;46(3):494-503. Epub 2004/12/18. doi: 10.1194/jlr.M400362-JLR200. PubMed PMID: 15604523.

12. Pramme-Steinwachs I, Jastroch M, Ussar S. Extracellular calcium modulates brown adipocyte differentiation and identity. *Sci Rep.* 2017;7(1):8888. Epub 2017/08/23. doi: 10.1038/s41598-017-09025-3. PubMed PMID: 28827782; PubMed Central PMCID: PMCPMC5567186.

13. Bae YK, Kwon JH, Kim M, Kim GH, Choi SJ, Oh W, et al. Intracellular Calcium Determines the Adipogenic Differentiation Potential of Human Umbilical Cord Blood-Derived Mesenchymal Stem Cells via the Wnt5a/beta-Catenin Signaling Pathway. *Stem Cells Int.* 2018;2018:6545071. Epub 2018/08/21. doi: 10.1155/2018/6545071. PubMed PMID: 30123291; PubMed Central PMCID: PMCPMC6079381.

14. Christodoulides C, Lagathu C, Sethi JK, Vidal-Puig A. Adipogenesis and WNT signalling. *Trends Endocrinol Metab.* 2009;20(1):16-24. Epub 2008/11/15. doi: 10.1016/j.tem.2008.09.002. PubMed PMID: 19008118; PubMed Central PMCID: PMCPMC4304002.

15. Song HI, Yoon MS. PLD1 regulates adipogenic differentiation through mTOR - IRS-1 phosphorylation at serine 636/639. *Sci Rep.* 2016;6:36968. Epub 2016/11/23. doi: 10.1038/srep36968. PubMed PMID: 27872488; PubMed Central PMCID: PMC5181839.
16. Faust GG, Hall IM. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics.* 2014;30(17):2503-5. Epub 2014/05/09. doi: 10.1093/bioinformatics/btu314. PubMed PMID: 24812344; PubMed Central PMCID: PMC4147885.
17. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-303. Epub 2010/07/21. doi: 10.1101/gr.107524.110. PubMed PMID: 20644199; PubMed Central PMCID: PMC2928508.
18. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods.* 2018;15(8):591-4. Epub 2018/07/18. doi: 10.1038/s41592-018-0051-x. PubMed PMID: 30013048.
19. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122. Epub 2016/06/09. doi: 10.1186/s13059-016-0974-4. PubMed PMID: 27268795; PubMed Central PMCID: PMC4893825.
20. Diaz-Gay M, Vila-Casadesus M, Franch-Exposito S, Hernandez-Illan E, Lozano JJ, Castellvi-Bel S. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. *BMC Bioinformatics.* 2018;19(1):224. Epub 2018/06/15. doi: 10.1186/s12859-018-2234-y. PubMed PMID: 29898651; PubMed Central PMCID: PMC6001047.

21. Dees ND, Zhang Q, Kandoth C, Wendl MC, Schierding W, Koboldt DC, et al. MuSiC: identifying mutational significance in cancer genomes. *Genome Res.* 2012;22(8):1589-98. Epub 2012/07/05. doi: 10.1101/gr.134635.111. PubMed PMID: 22759861; PubMed Central PMCID: PMC3409272.
22. Wendl MC, Wallis JW, Lin L, Kandoth C, Mardis ER, Wilson RK, et al. PathScan: a tool for discerning mutational significance in groups of putative cancer genes. *Bioinformatics.* 2011;27(12):1595-602. Epub 2011/04/19. doi: 10.1093/bioinformatics/btr193. PubMed PMID: 21498403; PubMed Central PMCID: PMC3106187.

Figure Legends

Fig. 1 Variants identified in ten lipomas using whole exome sequencing. (A) Total number of variants found in each lipoma sample. (B) Number of each type of variant identified in ten lipoma samples.

Fig. 2 Signature of mutational process in lipomas. (A) Mutational signatures presented according to 96-substitution classification highlighting signature 15 as major process associated with lipomas. (B) Histogram shows percentage of signatures present in each lipoma sample.

Supplementary Fig. 1: Mutation signature associated with each lipoma patient.

Supplementary Fig. 2: Enriched pathways (A) and biological processes (B) identified and associated with validated mutated gene products in lipoma tumors.

A **B**

Figure 1

Figure 2