의학박사 학위논문

Understanding Genetic Drivers and Evolution Pathway of Seborrheic Keratosis

지루각화증의 유전학적 발생과 진행과정 기작의 이해

2019 년 1 월

서울대학교 대학원

의과학과 의학 전공

이 모 제

Understanding Genetic Drivers and Evolution Pathway of Seborrheic Keratosis

지루각화증의 유전학적 발생과 진행과정 기작의 이해

by
Moses Lee

A thesis submitted to the Department of Biomedical Sciences in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Medical Science at Seoul National University College of Medicine

January 2019

Approved by Thesis Committee:

Professor _____Chairman
Professor _____Vice chairman
Professor
Professor
Professor

Understanding Genetic Drivers and Evolution Pathway of Seborrheic Keratosis

지루각화증의 유전학적 발생과 진행과정 기작의 이해

Moses Lee

Department of Biomedical Sciences

Medical science

Seoul National University College of Medicine

Same genetic modifications are shared by both benign and malignant tumors. Understanding this different cellular, histologic and clinical outcome could reveal new strategy in cancer treatment. However, genetic landscape of benign hyperplasia is largely understudied. Seborrheic keratosis (SK) is a benign skin hyperplasia with no known related malignancy. *FGFR3* and *PIK3CA* are known drivers of SK, but no whole-genome level study has been done, and reason for its senescence is not fully revealed. In this study, 51 SK samples from 49 patients were studied in the whole-genome level. Somatic mutation, copy number alteration (CNA), and loss of heterozygosity (LOH)

events were found from 8 pilot samples using whole exome sequencing (WES), and recurrent driver mutation were confirmed with Sanger sequencing and SNP arrays in 43 follow-up samples. *FGFR3*, *ZNF750*, and *PIK3CA* were identified as key drivers of SK. Other potential drivers include *NOTCH1*, chr3q29 microdeletion, *MSH2-MSH6*, *TERT1* promoter, and *RB1*. *FGFR3* with *ZNF750* LoF only occurs in the acanthotic subtype and is sufficient to induce SK without any more genetic change. *FGFR3* with *PIK3CA* induces hyperkeratotic subtype of SK. The order of genetic events and clonal evolution pathway was determined by minor allele frequency analysis and laser microdissection. SK somatic mutation profile revealed the UV effect. Thus, SK shares many genetic pathways with other malignant events, and its genomic landscape provides various level of complexity in its development.

Keywords: Whole exome sequencing, Seborrheic keratosis, *FGFR3*, *ZNF750*, *PIK3CA*, Tumor evolution

Student Number: 2012-22262

Table of Contents

List of Tables and Figures

Tables

Figures

Introduction

1. Cancer genomics using NGS

The next-generation sequencing (NGS) and its applications are considered as the most important tools of molecular and genetic research fields, after its first application on rare disease (1). One of the other important applications was made on cancer. Through NGS, genetic foundation of each cancer is rapidly establishing. Researchers put international efforts to collect cancer samples, such as The Cancer Genome Atlas (TCGA) and The International Cancer Genome Consortium (ICGC). As cancer samples and whole-exome, whole-genome, RNA sequencing, and other datasets were cumulated, more and more oncogenes and tumor suppressor genes are discovered (2-6). Driver genes are found by recurrent discovery that cannot be explained statistically. Oncogenes tend to have a few mutational hotspots, while tumor suppressor genes show recurrent loss of function (LOF). Based on statistics, classification of the precise roles of each genetic alteration in cancer population is ongoing. About 140 cancer drivers are now confirmed, grouped into three main functional categories: cell fate, cell survival and genome maintenance (7). The basic strategy for discovering genetic foundation of cancer is schematically showed in Figure 1A.

All tumors are not the same, in their genetic cause, histology, and clinical outcome. There are notoriously malignant cases, such as pancreatic adenocarcinoma, small cell lung cancer, as well as benign tumors without any malignant potential (8). Each kinds of tumor show its different nature in molecular and clinical history as well. Some cancers such as colon cancer development take decades. It is preceded by precancerous polyp, and every level in its evolutional history shows step-by-step, progressive accumulation of carcinogenic change in oncogenes such as *APC*, *CTNNB1*, *KRAS*, *BRAF*, *PIK3CA*, etc. (9, 10). Simulations based on mutation rates and cancer prevalence predicts that at least 3 or more consequent genomic change that favors carcinogenesis should occur (11). On the other hand, some types of malignancies could be driven by a single genetic event. For example, chronic myeloid leukemia could be driven by single translocation event, *BCR-ABL* (12, 13). Based on tumor genetic information, clinical trials on targeted anticancer drugs are proposed. Clinicians use sequencing information of a patient to select most effective drug, opening the era of individual targeted therapy.

Recent development in acquisition and sequencing of small amount of cell (<10000) enabled the temporal reconstruction of mutational lineage of a single cancer mass (14). Sequence of mutation could also be evaluated by bulk sequencing and estimating clonal prevalence of the cells based on mutation minor-allele frequency and copy-number alteration, but insufficient

in accurate temporal prediction of mutational sequence (15, 16). Information of sampling site and mutational landscape of individual resections showed that a cancer mass is a union of heterogeneous cells, and its genetic lineage could be precisely estimated in colorectal and esophageal cancers. Clonal structure revealed that driver somatic mutation and copy number variation (CNV) are shared by whole population, while passenger mutations are less prevalent, found only in small region (14, 17).

Based on the driver and passenger mutations each clones harbor, speed of cellular process of division, invasion, and metastasis differs, and so does its clinical and histologic feature.

2. Malignancies of the skin

Skin can be divided into 3 layers; epidermis, dermis, and subcutaneous fat layers. Each layer has different kinds of cells and thus gives rise to different benign and malignant neoplastic lesion. The normal epidermis of the skin have 2 type of cells: keratinocytes and melanocytes. Keratinocytes give rise to non-melanoma skin cancer (NMSC), and benign hyperplasia such as SK. Melanocyte give rise to lesions such as malignant melanoma and benign epidermal nevi. 90% or more skin cancer is NMSC (18, 19). The most important risk factors of skin NMSC is ultraviolet radiation,

NMSC is pathologically and clinically divided into two types of skin cancer: basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). 80~85% of the NMSC is BCC. The mortality of BCC is very low because it rarely metastasizes into other organs, while SCC is more aggressive and frequently invades into other tissues, capable of causing death (18, 20). Exposure to ultraviolet radiation (UVR) is the most important exogenous risk factor of both melanoma and NMSC. Skin cancer incidence is higher in sunlight-exposed skin area, lighter skin-colored individuals, history and frequency of sunburn. On the other hand, sun protection behavior have protective effect (20). The effect of UV The proportion of each type of base change could be used as the sunlight exposure marker, due to specific C to T base change by UVB (21-23). Sunlight-exposed samples harbored more than 60% of single nucleotide base change was C to T, compared to 30~40% in lung squamous cell carcinoma (24).

Sequencing of skin cancers revealed related cancer driver mutations in SCC (24) and BCC (22). Driver mutations of SCC were *TP53*, *CDKN2A*, *NOTCH2*, *NOTCH1*, etc. CNVs found were duplication in chromosomes 7, 8q, 9q, 14, and 20, and deletions at 3p, 4, 5q, 8p, 9p, 11, 17p, 18, 19, and 21(24). BCC driver mutations include *PTCH1*, *TP53*, *DPP10*, etc. *PTCH1* mutation was associated with LOH or deletion in chromosome 9, resulting in complete loss of functional protein. *TP53* was not associated with copy loss (22). Driver mutations in melanoma are *TP53*, *CDKN2A*, *NF1*, *RB1*, *ARID2*,

CBL, *BRAF*, *ERBB2*, *and EGFR* etc. MAP kinase, p53, RB1 pathway genes are widely involved in driving melanoma (23).

3. Clinical features of SK

Seborrheic keratosis (SK) is an extremely benign tumor that almost never develops into malignancy, only to be removed by cosmetic reasons (25-27). Sunlight exposure and increasing age are well known independent risk factors. SK are not observed until three or fourth decades of life, but its prevalence reaches 80~100% after fifth decade, showing extremely wide prevalence in geriatric population. It appears more frequently in sunlight-exposed parts of the body (face, forearm, dorsum of hand) than partial or non-exposed area (trunk, arms, inner forearms, legs), adjusted to body surface area (28).

Pathologically, SK shows keratinocyte hyperplasia involving all layers of epidermis. There are three main subtypes of SK, acanthotic, hyperkeratotic and adenoid subtypes, based on their degree of acanthosis, hyperkeratosis, papillomatosis, and pseudohorn cysts. Other subtypes include clonal, bowenoid, irritative, and melanoacanthoma subtypes (25). Naked eye and histologic view of two most frequent subtypes, acanthotic and hyperkeratotic SKs are shown in figure 1.

SK seems to be related with *AKT* activity, which blocks *TP53* mediated cell apoptosis by activation of *FOXO3*, and blocking it induces primary cultured SK cell death (29).

4. Genetic features of SK

Genetic causes of SK were first proposed in 2005. It was discovered that germline activating mutation S249C of *FGFR3* in mice shows benign epithelial neoplasm with histologic characteristics of acanthosis, hyperkeratosis, papillomatosis and keratin cyst resembling that of human SK. Known activating mutation sites of *FGFR3* were sequenced, and 40% of SKs showed *FGFR3* activating mutation (30). Variety of *FGFR3* mutations are found in inter and intra-individual SKs, emphasizing its crucial role in SK pathogenesis (31). Studies afterwards showed that majorities of seborrheic keratosis had *FGFR3* mutation, and at least some portion of it harbors *PIK3CA* mutations (27, 32). The mutations found in seborrheic keratosis are known to be associated with tumorigenesis, and same genetic changes are found in other malignant cancers.

FGFR3 is a tyrosine kinase and growth factor receptor. Germline mutation of *FGFR3* is known to be associated with achondroplasia, Crouzon syndrome, thanatotropic dysplasia I, II, and SADDAN syndrome, etc. (31). It

is also mutated in malignancies of bladder, lung, cervix, and multiple myeloma. K652M and R248C variants were known to activate *FGFR3* tyrosine kinase signal, by altering the position of regulatory domain (33) and facilitating dimerization of two *FGFR3* molecules, respectively (34). *FGFR3* has 3 major isoforms, FGFRIIIc, FGFRIIIb, FGFR3deltaTM. FGFRIIIc is the canonical isoform of *FGFR3*, but is expressed mainly on brain. FGFRIIIb isoform is mainly expressed in skin, while FGFR3deltaTM is expressed in both (35). In this study, *FGFR3* indicates FGFRIIIb transcript/protein. *FGFR3* activating K652M and R248C mutations are frequently found in normal human skin as well, but as a sole proliferative mutation, its tumorigenic effect is not enough, causing at most mild hyperplasia (34, 36). Also, no previous reports have found *FGFR3* participating as major tumor driver of skin malignancies such as squamous cell carcinoma, basal cell carcinoma, and cutaneous melanoma (22-24).

PIK3CA gene encodes catalytic subunit of an enzyme phosphoinositide 3-kinase (PI3K), which activates downstream mTOR, AKT, and MAPK pathway in response to ligand activated growth factor receptor (37, 38). Early embryonic somatic mosaicism of *PIK3CA* leads to relatively benign overgrowth of body parts. Mutations in *PIK3CA* is crucial in its tumorigenesis of breast, colon, endometrial, ovarian and gastric cancer (38, 39), and it has potential to change a benign breast cancer cell to be cancerous without any other mutation (40). PIK3CA mutation was rarely activated in

skin malignancies or pre-malignant lesion keratoacanthoma (23, 41). It is not found in basal cancer carcinoma or aggressive cutaneous squamous cell carcinoma (22, 24).

SK developmental process is schematically drawn in Figure 1B.

5. Why study genomics of SK?

Despite remarkable genetic findings in SK, some questions and missing link still remains.

First, benign property of seborrheic keratosis does not seem to go along well with its aggressive mutational profile: *FGFR3* and *PIK3CA*. Not a single invasive skin lesion has been proven to arise from SK, even with high incidence of *FGFR3* and *PIK3CA*. Some evidence on its benign nature has been suggested based on observation of *FGFR3*/*FOXN1* positive feedback loop. When *FGFR3* was activated, transcription factor *FOXN1* was transcribed, which in turn induces keratin expression and differentiation of keratinocyte (42). However, the result is still confounding because there are studies that both succeed (29) and failed (43) in observing *FOXN1* upregulation, which used qPCR and microarray, respectively. On top of these discrepancies, the *in vitro* system utilized for studying expression of *FOXN1* was cutaneous carcinoma cell line, which cannot represent normal

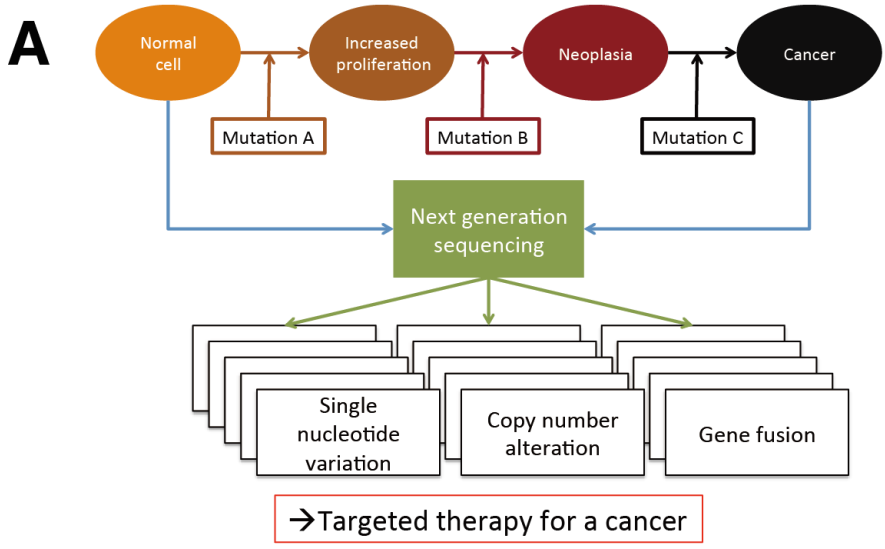
keratinocyte behavior. *FGFR3* activation mutation failed to induce *FOXN1* expression in normal skin xenograft sample (34). To summarize, no suitable explanation for the benign property of SK have been suggested.

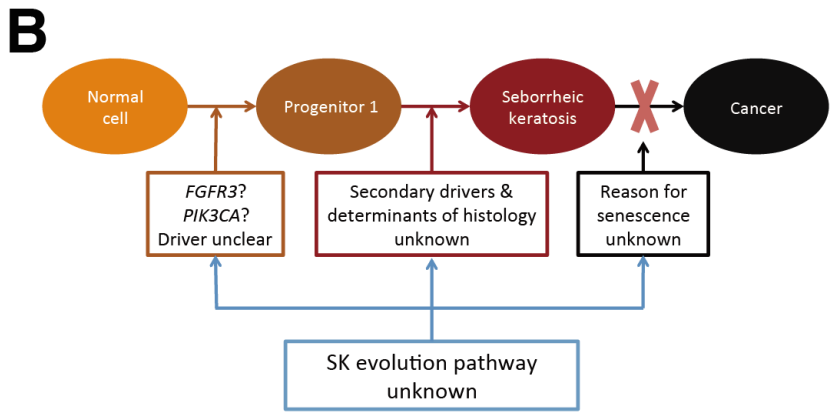
Second, the proliferative ability of *FGFR3* or *PIK3CA* to drive SK is unclear. Although the genes are known as notorious oncogenes, previous reports questions their oncogenic potential in skin. For instance, a WES study searching for somatic mutation accumulation in normal skin revealed *FGFR3* activation mutation in oncogenic site. The study utilized normal eyelid skin sample from a patient who removed it for treatment of dermatochalasis (36). Proliferation study for Also, *FGFR3* and *PIK3CA* are not found as major oncogene in other epithelial cell malignancies (22-24).

Lastly, the interaction between two previously known drivers is not assessed beforehand. Interactions between tumor drivers are reported many times beforehand. If two oncogenes inhibit proliferative effect of each other or participate in same pathway, they tend to segregate in different clone. For instance, *TP53* and *MDM2* in bladder cell carcinoma have mutually exclusive distribution (44). Looking for possible interaction could reveal mechanism for SK senescence. (Figure 1C.)

In this study, I tried to identify the causal mutations of SK. 51 SK samples from 49 patients were initially cut for histologic confirmation. With leftover sample, DNA was extracted. I first whole-exome sequenced (WES) 8 SK samples with its normal salivary counterparts from 7 patients who

removed the tissues for cosmetic reasons. Based on WES data, I identified causal genes of SK, *FGFR3*, *ZNF50* and *PIK3CA*, and they were statistically significant after searching 43 other samples of SK with Sanger sequencing (Figure 1D). Surprisingly, *ZNF750* and *PIK3CA* mutations segregated based on SK subtypes, along with *FGFR3* activating mutations. After confirming driver mutation in SK, I evaluated the evolution and mutational sequence of SK via minor allele frequency and laser microdissection.


A
Normal cell
Increased proliferation
Neoplasia
Cancer
Mutation A
Mutation B
Mutation C
Next generation sequencing
Single nucleotide variation
Copy number alteration
Gene fusion
→Targeted therapy for a cancer


B
Normal cell
Progenitor 1
Seborrheic keratosis
Cancer
FGFR3? PIK3CA? Driver unclear
Secondary drivers & determinants of histology unknown
Reason for senescence unknown
SK evolution pathway unknown

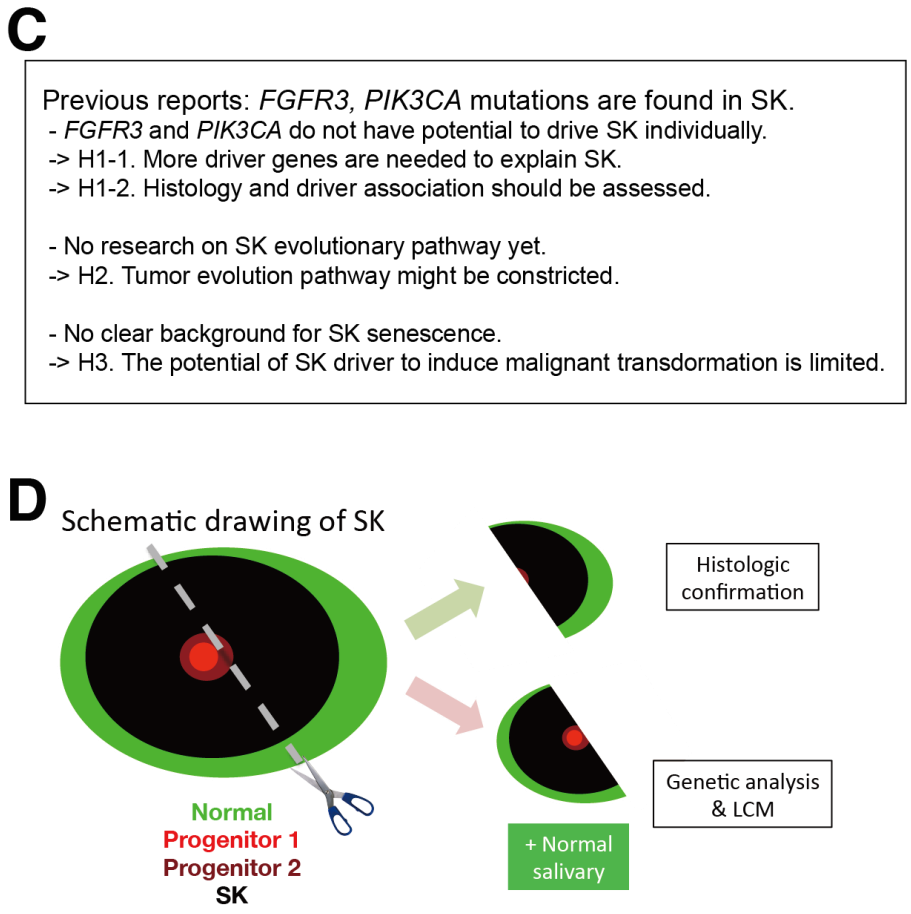


Figure 1. Introduction of cancer genomics, previous findings of SK, hypothesis and experiment design. A. The schematic drawing for cancer genomics. Multiple cancer samples are analyzed by next generation sequencing in order to determine their drivers. When clustered in large number, genomic landscape of certain cancer type could be determined. **B.** The known driving stages of SK. *FGFR3* and *PIK3CA* are known as SK

drivers, but their exact role, evolutionary pathway, and reason for senescence is yet unrevealed. **C.** The key question and hypothesis of this study. **D.** Schematic drawing of experiment design. 51 SK samples are initially cut in half, each piece used for histologic confirmation and genetic analysis.

Materials and Methods

Tissue Acquisition

All tissues used in this study were acquired from the patient who removed the legion for cosmetic reason. The patients were verbally informed and voluntarily signed to the informed consent. Genetic study was approved from in-hospital IRB from Samsung Medical Center.

Whole exome sequencing and variant calling

The procedures used for the preparation of genomic DNA, whole exome capture using an Agilent V5 array, sequencing using an Illumina HiSeq 2500 platform (San Diego, CA, USA). Sequenced reads were aligned to humen genome 19 (hg19) with ELAND program. Reads outside human genome are discarded. The variants are called based on proportion of non-hg19 base. The probability of variant called by random error follows binomial distribution with error rate given by manufacturer. P-value for each variant was calculated based baysian statistics based on error rate and coverage. Variants that passed p-value were assessed through gene information provided by UCSC, filtered by assessed protein modifying potential. Normal population based filtering through 1000 genome, Exome Aggregation Consortium was done to select rare variants. Somatic variants and CNA/LOH are called based on comparison

between SK sample and normal salivary counterpart of each patient. All processes of WES and somatic variant calling have been described in detail previously (45).

Sanger sequencing

PCR amplification was performed with 10 pmol of each specific primers. The PCR conditions were an initial denaturation at 95 °C for 3 min, followed by 35 cycles of amplification (95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s) and a final extension at 72 °C for 5 min. The PCR product was gel extracted and Sanger sequenced on an ABI3730XL DNA Analyzer instrument (Applied Biosystems, Foster City, CA, USA).

The primer for *FGFR3* is forward 5'-CGTGGAGAACAAGTTTGGCA - 3' and reverse 5' – CTCCAACCCCTAGACCCAAA - 3' for R248C and R249C, forward 5'-TCCCACATCCTGCCTCGTGC - 3' and reverse 5' – CCCGCTCCGACACATTGGC - 3' for S311F, forward 5'-CAGTGTGTATGCAGGCATCC - 3' and reverse 5' - GGAGATCTTGTGCACGGTGG - 3' for Y375C, A393E, forward 5' - GCTGGTGACCGAGGACAACG - 3' and reverse 5' - GGTGTTGGCGCCAGGCGTCC – 3' for K652M.

The primer for *ZNF750* is forward 5'-ATGACGGAGGCTCTCGCCAGAC - 3' and reverse 5' - TGTTTTTTTGAGAAGCAGCAGCTGC - 3' for whole gene amplification, read with 4 individual reading primer sets 5'-AAGCAAAGCGGGAGGTGCCTC - 3', 5'- CCACTGCCAAGGCCGTGTC - 3', 5' - CCATCTCCAGCCACATACGA - 3', and 5'-GCCTTCCCCGGTCGACCAC - 3'. Individual primer sets for each *ZNF750* variants are forward 5'- TGAGTCTCCTCAAAGAGCGG - 3' and reverse 5' - TAATCGAGTTTTTACAAAGACC - 3' for P19fsQ29X, forward 5' - GCATATCTGTGTGGGGGTGC - 3' and reverse 5' – CACCTGCAGATATGTACACATGC - 3' for H43Y, forward 5'-TGGGTTCCGTAGACTGACAG - 3' and reverse 5' - CCGAGACACTGGCTTTACAC - 3' for E217X, and forward 5'-GAAGGACTCGAGGCTGGATA - 3' and reverse 5' – TCCCTCTAACCTGCCGATTC - 3' for W343X.

The primer for PIK3CA is forward 5'-AATGGGGAAAAAGGAAAGAATG - 3' and reverse 5' - CCAACTCTAAGCATGGAGTTTCC - 3' for E453Q, forward 5'-CCAGAGGGGAAAAATATGACA - 3' and reverse 5' - CATTTTAGCACTTACCTGTGAC - 3' for E542K and E545K, forward 5'-CATTTGCTCCAAACTGACCA - 3' and reverse 5' - TGAGCTTTCATTTTCTCAGTTATCTTTTC - 3' for H1047R and H1047L.

SNP array

1ug of DNA was extracted from each tumor tissues using Geneall DNA extraction kit, According to the manufacturer's protocol. DNA fragmentation was checked through gel electrophoresis.

SNP array was done using Illumina HumanOmniExpress-24v1.1, according to the manual provided by the manufacturer. The result provides two parameters, B allele frequency and Log R ratio. CNV was called based on PennCNV protocol (46-48). LOH were called based on assumption that B-allele frequency and Log R ratio follows standard distribution in 100Mbps window, based on in-house variant calling script.

Laser microdissection

SK tissues were cut, positioned, and embedded in OCT compound inside properly sized cryomold. The samples were flash-frozen on the surface of liquid nitrogen. OCT-embedded samples were incubated for 5 minutes inside pre-cooled cryosection device, Leica CM1510-3 Cryostat, cryosected in 12-micrometer thickness in –18'C.

Cryosections were laid on MMI MembraneSlide, and H&E staining was done. Sample was initially fixed with 95% ethanol, air dried for 20 min. Hematoxylin was applied for 30 seconds, washed with RNase-free water for

30 seconds. Eosin applied for 30 seconds, and serial ethanol dilution of 70%, 95%, 100% were applied 15 sec each.

After H&E staining, laser microdissection was done by MMI CapLift laser microdissection system. Software instruction and laser control, focusing on sample, target selection is all done according to manufacturer's protocol. DNA from each microdissected samples was extracted with QIAamp DNA Micro Kit, according to the manufacturer's protocol.

Chapter 1

Genetics of acanthotic SK

Results

1. WES of SK reveals causal genes of acanthotic SK: *FGFR3*, and *ZNF750* & other candidates

The pilot study included 8 samples, 5 acanthotic and 3 hyperkeratotic SK (Figure 2), randomly selected from 51 SK samples (Figure 3A). Total 878 somatic mutations were found in all of the samples, 611 of them protein altering, 50 damaging, and 33 in known Catalogue of Somatic Mutations In Cancer (COSMIC) position (Figure 3B). 3 of the samples were hyperkeratotic subtypes, and 5 were acanthotic subtypes (Figure 3C), confirmed with pathologic diagnosis. All SK samples were covered at least 120x and normal salivary samples were covered at least 60x (Table 1) and LOH/CNV search based on coverage information revealed recurrent LOH at chr17q site (Figure 3D, 4). Through in-house somatic variant calling pipeline, I identified total 293 mutational events from 5 acanthotic subtypes. Histologic information of all sequenced sample are presented at table 2.

Genes with recurrent mutations over 2 or more samples are *FGFR3, ZNF750, ABCA4, SYNE1, and GFRAL*. Of these, 3 passed Bonferroni correction in binomial test result: *FGFR3, ZNF750,* and *GFRAL* (Table 3). Filtering through population database (1000Genomes, Exome Aggregation consortium (ExAC)), COSMIC, and in-house database, I concluded that 2

genes are profound candidates for SK acanthotic subtypes: *FGFR3* and *ZNF750*.

FGFR3 variant was found in all 5 samples, reaching genome-wide statistical significance. Only one variant was repeatedly observed in all samples, K652M (Table 4). K652M variant is known to constitutively activate *FGFR3* kinase (33). However, in epidermal area, single activation mutation of *FGFR3* is not strong enough to cause hyperplasia; skin with normal appearance was found to harbor *FGFR3* R248C mutation (34, 36). *ZNF750* variants were destructive in its nature, showing frameshift (P19QfsTer26), nonsense (Q29X), and missense mutation in conserved domain (H43Y). Moreover, all *ZNF750* mutated samples harbored large LOH in 17q region (Figure 2D, 3, Table 4), eradicating WT copy of *ZNF750*. The series of genetic events might completely knockout wild type *ZNF750* expression in the samples. No knockout mutations of *ZNF750* were found in 1000 Genomes, Exome Aggregation Consortium (ExAC), and in-house Korean database.

ZNF750 is expressed in epithelial cells covering skin and gastrointestinal tracts, and it is known to participate in late keratinocyte differentiation (49-52). Single frameshift mutation of *ZNF750* is related with autosomal dominant seborrhea-like dermatitis, though it is not clear whether the condition accompanied LOH (53). *ZNF750* mutation is also found in

esophageal cancer, with strong association of LOH and bi-allelic inactivation (17, 54). This is the first report of *ZNF750* inactivation in SK. This loss of function and LOH pattern most likely resembles the mutational patterns of tumor suppresser genes (7).

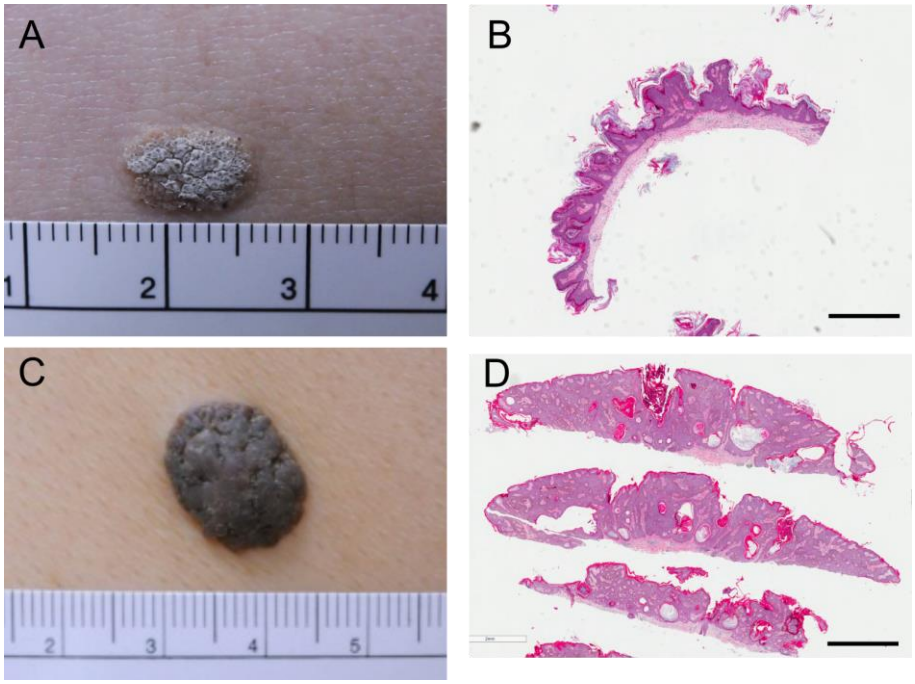


Figure 2. Histologic example of major SK subtypes. A. Hyperkeratotic SK seen with naked eye and **B.** its histology. **C.** Acanthotic SK seen with naked eye and **D.** its histology. Scale bar = 1mm. Skin picture & histology data provided and permitted by Jiho Park and Hae Yong Yoo, Samsung Advanced Institute for Health Sciences & Technology.

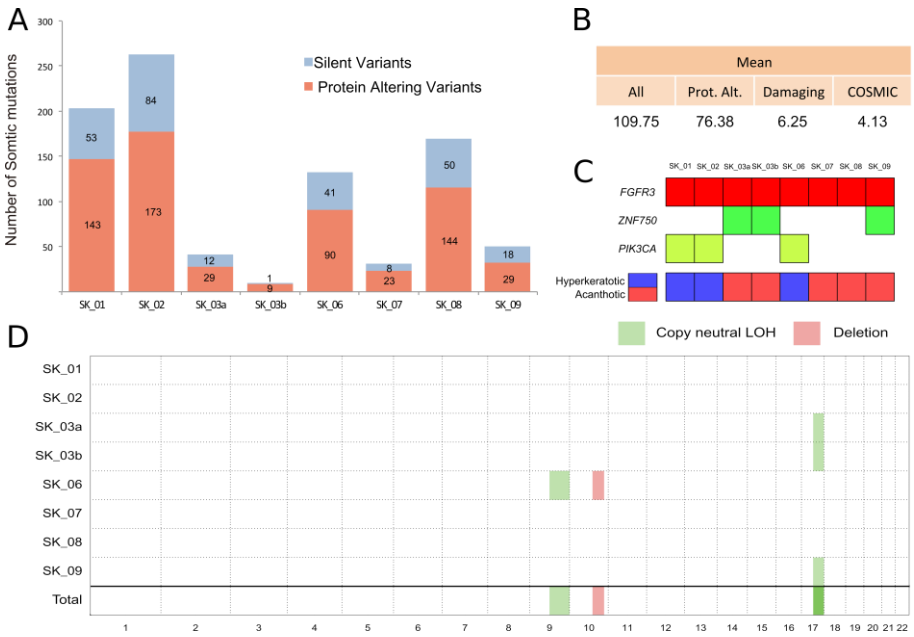


Figure 3. WES and CNV profile of 8 pilot SK samples. A. The number of somatic mutation found in each samples. **B.** The mean value of somatic mutation of each group. **C.** The distribution of *FGFR3*, *ZNF750*, *PIK3CA* with sample histologic subtypes. **D.** CNV and LOH profile of 8 SK samples. 3 samples had LOH at chr17q region, and SK_06 had LOH and deletion at chr9 and chr10, respectively.

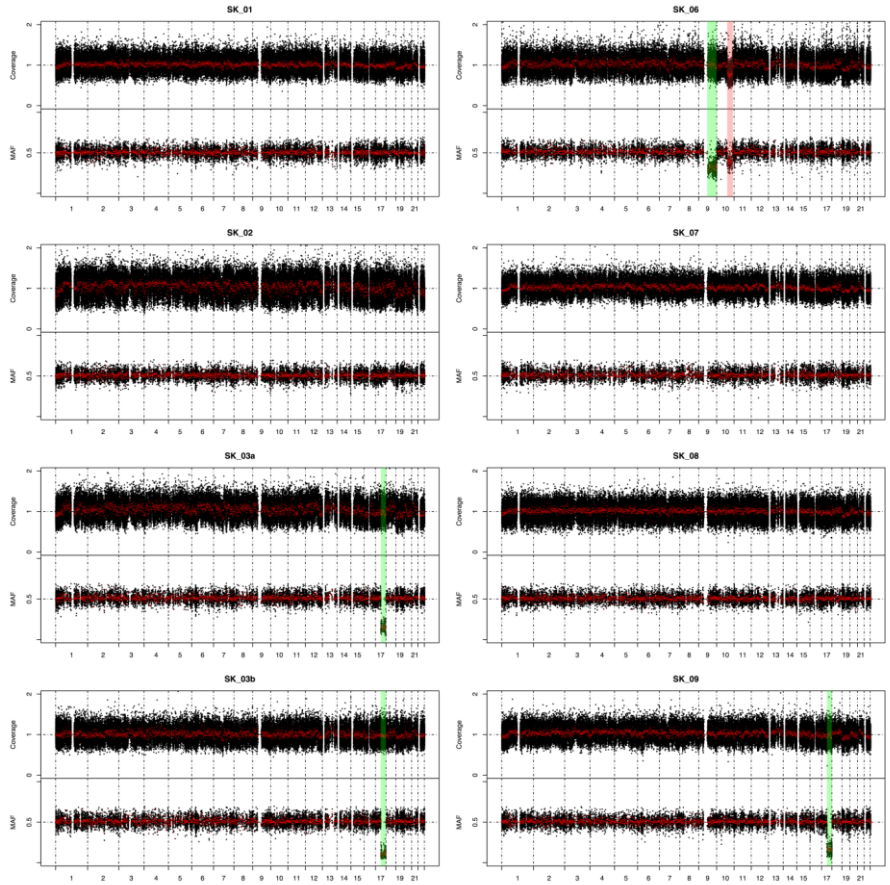


Figure 4. CNV and LOH called from WES coverage information. Red highlight indicates deleted region, green highlight indicates LOH region. SK03a, 03b, 09 have chr17q somatic LOH. SK06 shows chr10q deletion and chr9q LOH.

Table 1. Whole Exome Sequencing statistics of 8 SK samples.

Sample	NM_01*	SK_01	NM_02	SK_02	NM_03	SK_03a	SK_03b	NM_06	SK_06	NM_07	SK_07	NM_08	SK_08	NM_09	SK_09
Sex	Male	Male	Male	Male	Male	Male	Male	Female	Female	Male	Male	Male	Male	Male	Male
Tissue Type	Normal	Tumor	Normal	Tumor	Normal	Tumor	Tumor	Normal	Tumor	Normal	Tumor	Normal	Tumor	Normal	Tumor
Read length (bps)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Number of reads (M)	85.87	105.7	81.54	143.79	58.03	104.04	113.49	52.91	112.13	49.54	97.92	82.36	137.73	75.32	130.73
Mean coverage (x)	101.43	127.4	100.21	173.72	70.39	122.18	135.65	65.78	135.26	59.71	118.2	98.29	161.63	87.24	154.3
% of reads mapped on genome	98.79	99.4	98.52	99.44	98.87	99.52	99.6	98.46	99.57	99.26	99.58	97.83	99.44	97.33	99.38

* NM: normal counterpart sample.

% of reads mapped on target	74.88	76.51	77.81	77.15	76.43	74.79	76.26	77.6	76.75	76.82	76.82	75.81	74.86	73.99	75.21
% of targeted bases covered at least 4x	99.8	99.82	99.8	99.84	99.7	99.89	99.87	99.56	99.78	99.62	99.86	99.73	99.84	99.7	99.88
% of targeted bases covered at least 8x	99.48	99.55	99.49	99.62	99.08	99.7	99.66	98.88	99.62	98.81	99.65	99.34	99.62	99.24	99.7
% of targeted bases covered at least 20x	97.52	98.07	97.59	98.55	94	98.38	98.47	93.24	98.61	91.77	98.22	97.09	98.54	96.45	98.72
Mean error rate (%)	0.18	0.23	0.18	0.23	0.13	0.13	0.13	0.13	0.14	0.13	0.13	0.17	0.24	0.17	0.23

Table 2. The clinical information of WES samples.

Sample No.	Age/Sex	Histologic subtype	Location	No. of Somatic mutation	% of Tumor cells
SK01	73/M	Hyperkeratotic	Left cheek	220	49.6~70.3%
SK02	78/M	Hyperkeratotic	Left temple	279	74.5~82.6%
SK03a	44/M[†]	Acanthotic	Back	43	69.3~81.0%
SK03b	45/M[*]	Acanthotic	Belly	10	77.8~90.1%
SK06	72/F	Hyperkeratotic	Posterior, lower of Rt. ear	139	36.9~37.7%
SK07	66/M	Acanthotic	Hairline, posterior of left ear	36	77.3~87.5%
SK08	57/M	Acanthotic	Left scalp	187	42.3~50.5%
SK09	59/M	Acanthotic	Back	59	67.2~78.9%

[†] SK03a and 03b samples are biopsied from same individual in 2 different times.

Gene	Total # of Variants	Conserved‡	Silent	Protein Altering	Damaging§	Binomial-total	Binomial-altering	Binomial-conserved	Binomial-damaging
FGFR3	5	5	0	5	0	1.498.E-10‡	2.472.E-11‡	5.764.E-13‡	9.982.E-01
ZNF750	3	2	0	3	3	2.668.E-06‡	9.076.E-07‡	3.439.E-05‡	7.451.E-10‡
GFRAL	2	2	0	2	0	3.150.E-04	1.539.E-04‡	3.439.E-05‡	9.984.E-01
ABCA4	2	0	1	1	0	2.948.E-03	5.263.E-02	9.742.E-01	9.948.E-01
SYNE1	2	1	0	2	0	3.478.E-02	1.851.E-02	9.113.E-02	9.802.E-01

Table 3. List of recurrently mutated genes in acanthotic SK. Binomial tests assumed random mutation over targeted area.

‡ Amino acid sequence difference equal or less than 2 in 100-species alignment.

§ Framshift, splicing, or nonsense variants.

‡ Passed Bonferroni test.

Table 4. WES result of oncogenic candidates of acanthotic SK.

Sample ID	Gene	Chromosome: position (hg19)	Base change	Somatic mutation	Impact on protein	Amino acid change	Amino acid location / protein length	COSMIC gene/ position	# of species different from human/ # of species with ortholog	Coverage depth (alternative allele/total)	
										Tumor	Normal
SK_03a	*FGFR3*	Chr4:1807890	A>T	AA>AB	Missense	K652M	652/808	True/True	0/0	75/154 (48.7%)	0/103 (0%)
SK_03b	*FGFR3*	Chr4:1807890	A>T	AA>AB	Missense	K652M	652/808	True/True	0/0	76/185 (41.1%)	0/103 (0%)
SK_07	*FGFR3*	Chr4:1807890	A>T	AA>AB	Missense	K652M	652/808	True/True	0/0	84/192 (43.8%)	1/95 (1.1%)
SK_08	*FGFR3*	Chr4:1807890	A>T	AA>AB	Missense	K652M	652/808	True/True	0/0	48/190	0/136

										(25.2%)	(0%)
SK_09	*FGFR3*	Chr4:1807890	A>T	AA>AB	Missense	K652M	652/808	True/True	0/0	69/165 (41.8%)	0/131 (0%)
SK_03a	*ZNF750*	Chr17:80790274	G(-)	AA>BB	Frame-shift Deletion	P19QfsTer26	19/723	True/na	-/-	204/252 (81.0%)	0/95 (0%)
SK_03b	*ZNF750*	Chr17:80790246	G>A	AA>BB	Nonsense	Q29X	29/723	True/na	0/0	195/215 (90.7%)	0/124 (0%)
SK_09	*ZNF750*	Chr17:80790204	G>A	AA>BB	Missense	H43Y	43/723	True/na	0/0	233/298 (78.2%)	0/210 (0%)

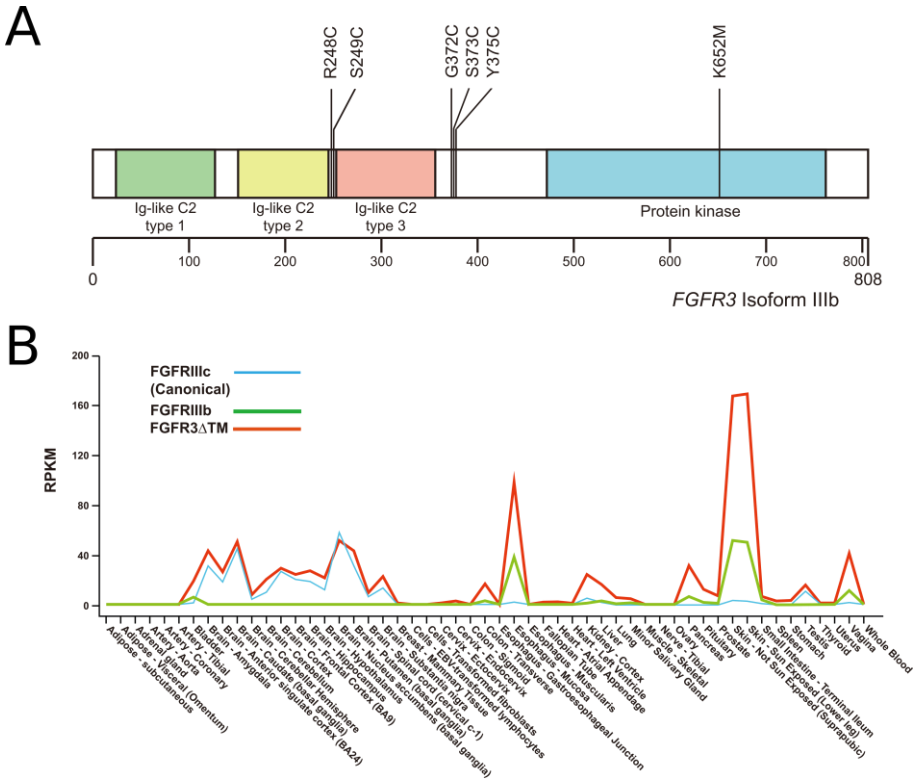
2. *FGFR3* and *ZNF750* in extended sample data sets

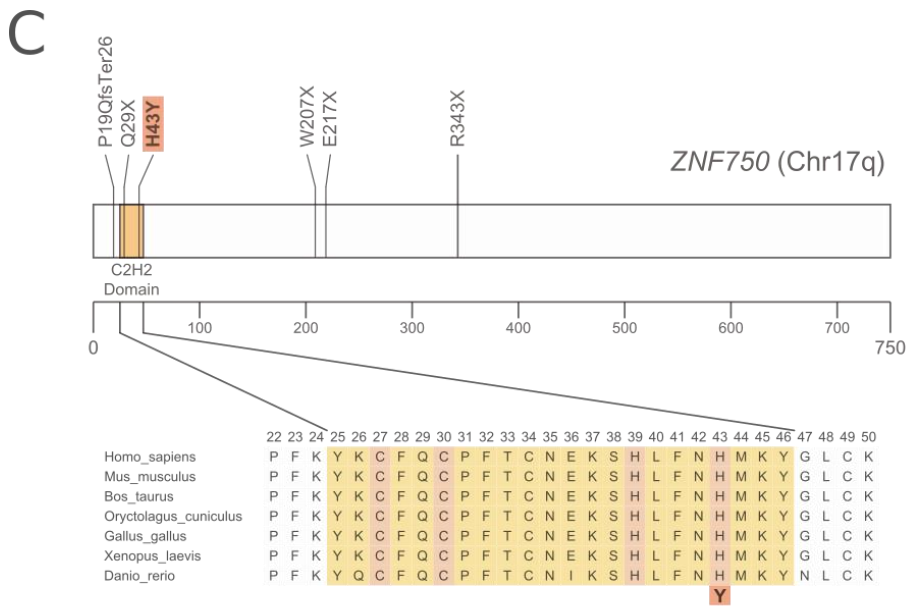
In order to replicate previously found *ZNF750* mutation, I Sanger sequenced all exons of *ZNF750* and *FGFR3* in WES samples and additional 30 pathologically confirmed acanthotic SK, with their normal salivary counterpart (Table 4).

23 of the 30 replication data set harbored *FGFR3* mutations. Seven of those were K652M mutation. Additional mutational spots included R248C (6 samples), S249C (7 samples), G372C (1 sample), S373C (1 sample), and Y375C (1 sample) (Figure 5A, Table 5). All of these mutations were previously reported to cause seborrheic keratosis (30). Mutations R248C, S249C, K652M is shared in both FGFR3deltaTM and FGFRIIIb isoforms, while G372C, S373C, and Y375C mutations are found only in FGFRIIIb isoform (Figure 5B).

Additional 4 more damaging nonsense mutations in *ZNF750* were found in replication data sets, passing genome-wide significance (Figure 5C, Table 6). In WES, I observed all of *ZNF750* mutations samples had LOH (Figure 3D), which is expected to complete eradicate *ZNF750* wild type protein expression. I could not confirm LOH event by Sanger sequencing because cancer cell proportion is not 100%. However, SK28, sample with *ZNF750* W207X had 2068bp microdeletion confined to *ZNF750* gene only (Figure 5D, Sanger sequencing data performed and permitted by Jiho Park

and Hae Yong Yoo, Samsung Advanced Institute for Health Sciences & Technology.). This finding adds another layer of confidence that *ZNF750* is crucial tumor suppressor gene in development of acanthotic SK.


A
R248C
S249C
G372C
S373C
Y375C
K652M
Ig-like C2 type 1
Ig-like C2 type 2
Ig-like C2 type 3
Protein kinase
0
100
200
300
400
500
600
700
800
808
FGFR3 Isoform IIIb
B
FGFRIIIc (Canonical)
FGFRIIIb
FGFR3ΔTM
RPKM
200
160
120
80
40
0
Adipose - Subcutaneous
Adrenal Gland
Artery - Aorta
Artery - Coronary
Artery - Tibial
Bladder
Brain - Amygdala
Brain - Anterior cingulate cortex (BA24)
Brain - Caudate (basal ganglia)
Brain - Cerebellar Hemisphere
Brain - Cerebellum
Brain - Cortex
Brain - Frontal Cortex (BA9)
Brain - Hippocampus
Brain - Hypothalamus
Brain - Nucleus accumbens (basal ganglia)
Brain - Putamen (basal ganglia)
Brain - Spinal cord (cervical c-1)
Brain - Substantia nigra
Breast - Mammary Tissue
Cells - EBV-transformed lymphocytes
Cells - Transformed fibroblasts
Cervix - Ectocervix
Cervix - Endocervix
Colon - Sigmoid
Colon - Transverse
Esophagus - Gastroesophageal Junction
Esophagus - Mucosa
Esophagus - Muscularis
Fallopian Tube
Heart - Atrial Appendage
Heart - Left Ventricle
Kidney - Cortex
Liver
Lung
Minor Salivary Gland
Muscle - Skeletal
Nerve - Tibial
Ovary
Pancreas
Pituitary
Prostate
Skin - Not Sun Exposed (Suprapubic)
Skin - Sun Exposed (Lower leg)
Small Intestine - Terminal Ileum
Spleen
Stomach
Testis
Thyroid
Uterus
Vagina
Whole Blood


C
P19QfsTer26
Q29X
H43Y
W207X
E217X
R343X
ZNF750 (Chr17q)
C2H2
Domain
0
100
200
300
400
500
600
700
750
22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
Homo_sapiens P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Mus_musculus P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Bos_taurus P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Oryctolagus_cuniculus P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Gallus_gallus P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Xenopus_laevis P F K Y K C F Q C P F T C N E K S H L F N H M K Y G L C K
Danio_rerio P F K Y Q C F Q C P F T C N I K S H L F N H M K Y N L C K
Y

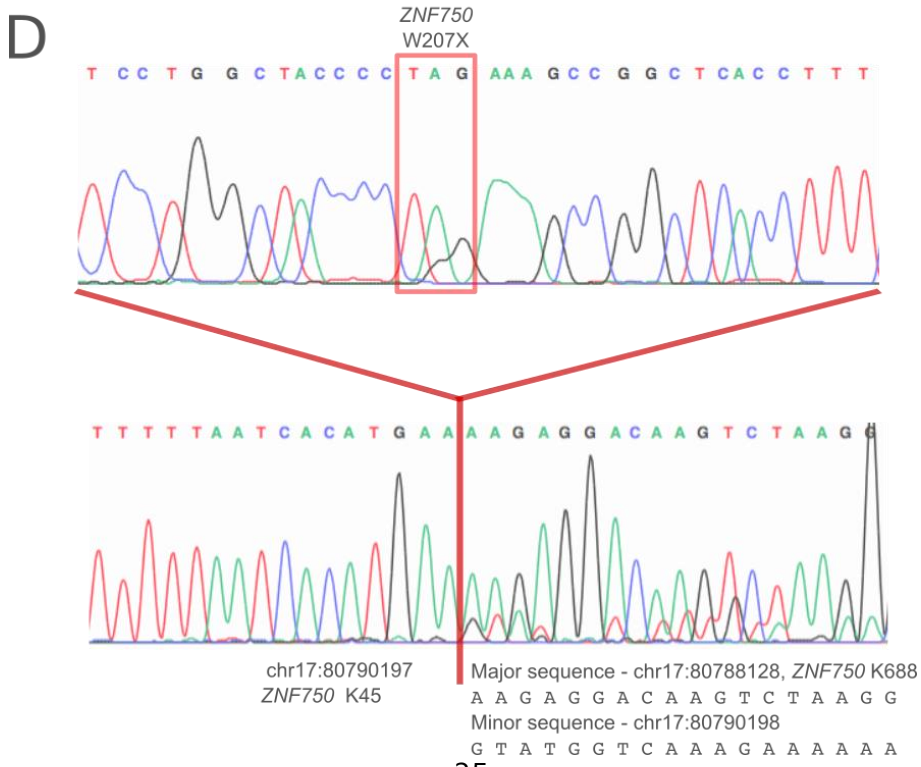

D
ZNF750
W207X
T CC T G G C TACCCC C TA G AAA GCC G G C T CAC C T T T
T T T T TAA T C A C A T G A A A A G A G G A C A A G T C T A A G G
chr17:80790197
ZNF750 K45
Major sequence - chr17:80788128, ZNF750 K688
A A G A G G A C A A G T C T A A G G
Minor sequence - chr17:80790198
G T A T G G T C A A A G A A A A A A

Figure 4. Cancer drivers found in WES of acanthotic samples. A. *FGFR3* gene structure and somatic mutation found in WES and Sanger sequencing of acanthotic SKs. **B.** Isoform of *FGFR3* expressed in each tissues. **C.** *ZNF750* mutations found in WES and Sanger sequencing of acanthotic SKs. **D.** Sanger sequencing example of *ZNF750* with smallest intragenic LOH size of 2068bps. Lys45 to Lys688 sequence with intron is deleted.

Table 5. Clinical information of the extended SK samples.

No.	Sex	Location	Type[5]	*ZNF750*	*FGFR3*	*PIK3CA*	SNP array
SK10	F	Back	A	R343X	K652M		O
SK11	F	Back	A		R248C		O
SK12	M	Back	A		R248C		O
SK20	M	Scalp	A		K652M		O
SK21	M	Back	A	E217X	K652M		O
SK22	M	Back	A	W207X	K652M		O
SK24	F	Thigh	A		S249C		O
SK25	F	Scalp	A		S249C		O
SK26	M	Penile glans	A		K652M		O
SK27	M	Rt. Groin	A		S373C		O
SK28	F	Lt. waist	A	W207X	S249C		
SK29	M	Lt. Scalp	A		S249C		O
SK30	M	Lt. Thigh	A		R248C		O
SK33	M	Face	A		K652M		O
SK34	M	Back	A				O
SK35	M	Thigh	A		R248C		O
SK36	M	Thigh	A		Y375C		O
SK37	M	Groin	A		R248C		O
SK38	F	Abdomen	A				O
SK40	M	Forehead	A				O
SK41	M	Back	A				O

[5] H: Hyperkeratotic, A: Acanthotic

SK43	M	Rt. Cheek	A	S249C	I45M-Germline	O
SK45	F	Lt. Inguinal	A			O
SK46	M	Lt. Scalp	A	G372C		O
SK47	F	Rt. Scalp	A	R248C	E453K	O
SK48	F	Lt. Scalp	A	K652M		O
SK49	M	Penile Shaft	A			O
SK50	M	Rt. Scalp	A	S249C	H1047L	
SK51	M	Lt. Scalp	A	S249C		
SK55	F	Rt. flank	A			
SK13	M	Abdomen	H	R248C	H1047R	O
SK14	M	Leg	H			O
SK15	F	Scalp	H			O
SK16	M	Scalp	H			O
SK19	M	Chest	H	G372C		O
SK23	F	Abdomen	H			O
SK31	M	Scalp	H			O
SK32	M	Abdomen	H			
SK39	F	Inguinal	H	Y375C	H1047R	O
SK42	F	Lt. shin	H			O
SK44	F	Back	H	K652E		O
SK53	M	Rt. Flank	H			

SK54	F	Lt. Shin	H		R248C	E545K + germline E542K

Table 6. P-value change of acanthotic driver *FGFR3* and *ZNF750* after Sanger sequencing 30 replication samples. Both genes passed genome-side significance level of 10^{-6}.

Gene	Binom-total	Binom-alter	Binom-con	Binom-damaging
FGFR3	6.127.E-50	2.575.E-54	1.890.E-63	3.988.E-83
ZNF750	9.379.E-10	7.852.E-11	4.285.E-13	5.341.E-18

3. CNV and LOH confirmed by SNP array

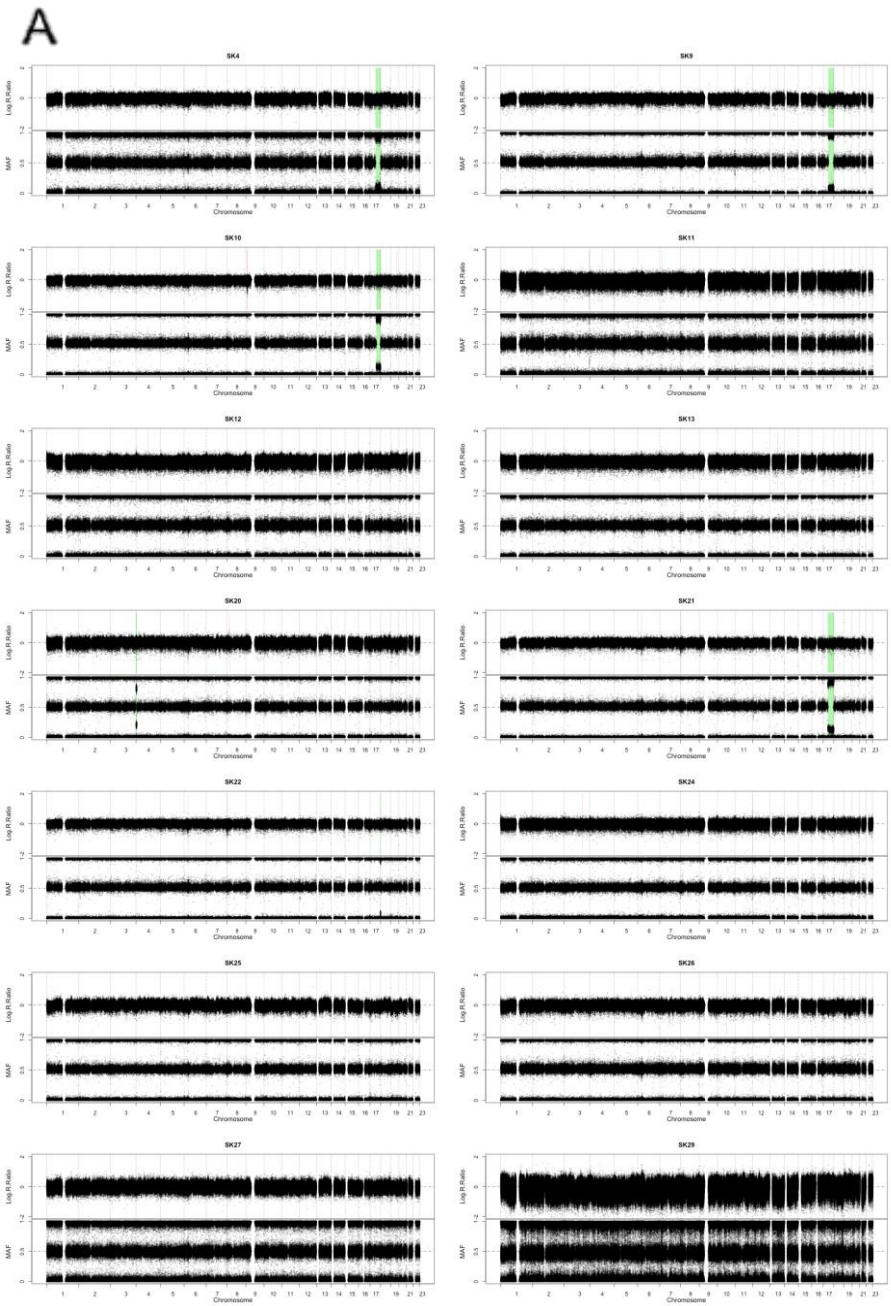
Through WES, I could detect CNV and LOH based on coverage and B allele frequency for 5 acanthotic SK samples. In order to confirm and search CNV and LOH of WES and replication acanthotic SK samples, 29 of 35 samples with available DNA were analyzed through SNP array (Table 5). SNP array data was trimmed, filtered, analyzed based on PennCNV tool (46-48). Each CNV were manually re-examined and visualized via in-house scripts (Figure 6, Table 7). LOH was called visualized, filtered, manually confirmed through in-house scripts.

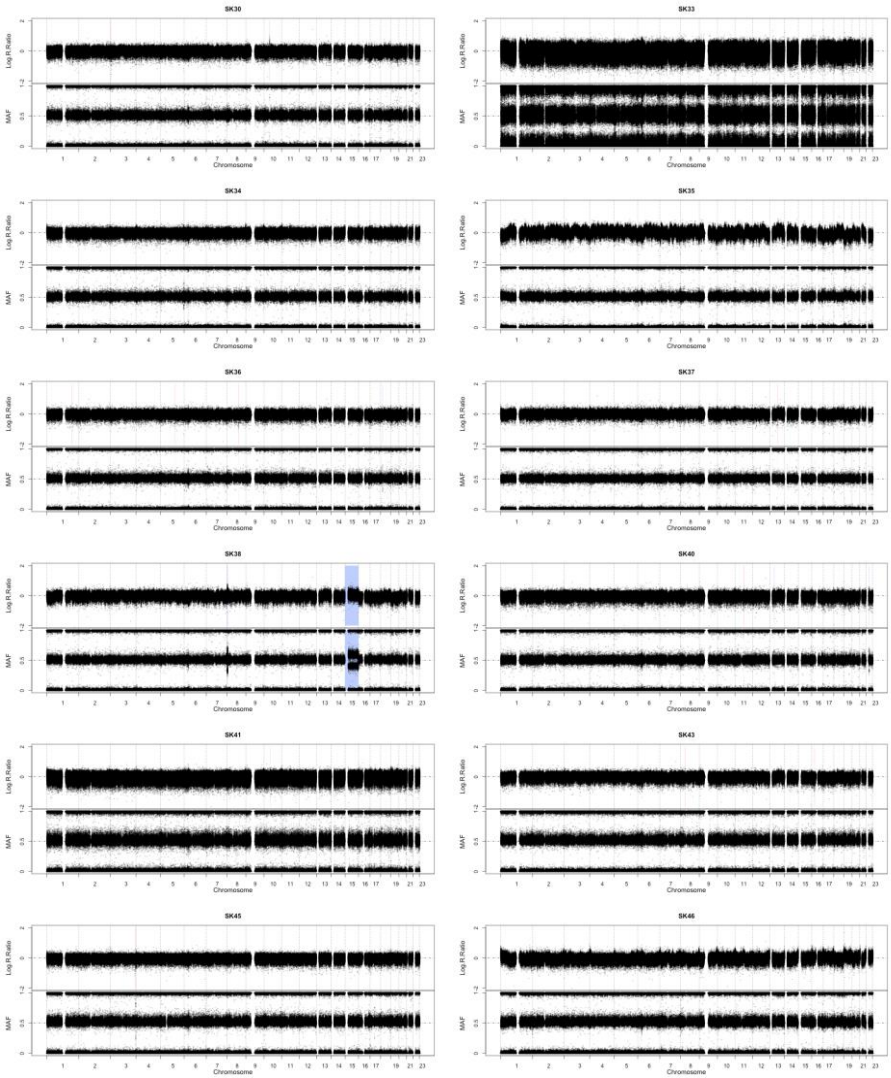
The result confirmed chr17 LOH in samples with *ZNF750* mutation, SK 03b and 09. New LOH in *ZNF750* was found in SK 10, 21, 22. Including 2068bps microdeleted SK28 sample, I found that all *ZNF750* loss of function mutations in acanthotic samples were associated with LOH event. Likewise, all LOH found in 17q, regardless of size, incorporates *ZNF750* locus and *ZNF750* mutation.

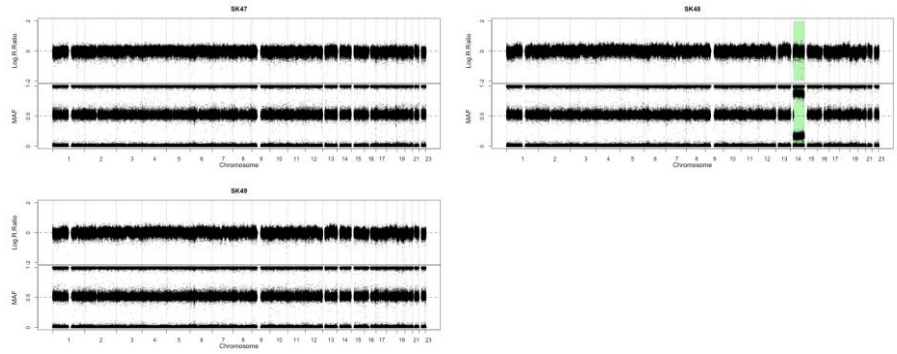
CNVs that were not found in DGV database (55), or normal population, was revealed by SNP array (Table 7). Two acanthotic samples, SK11 and 45 had 0.42 Mbps and 0.58Mpbs microdeleted in 3q29 region, and 0.38 Mbps were shared. The deleted area is known to be associated with 3q29 microdeletion syndrome (56). Major phenotypes in 3q29 microdeletion are

neural and facial defect mostly, but minor phenotype of abnormal skin pigmentation such as freckles, hairy nevus was also reported (57).

Other single sample large-scale CNV/LOHs include 6Mbps copy-neutral LOH in 4p, duplication of chr15, 8p. LOH in chr4p in SK20 includes *FGFR3* gene.


A
SK4
SK9
SK10
SK11
SK12
SK13
SK20
SK21
SK22
SK24
SK25
SK26
SK27
SK29
Log R Ratio
MAF
Chromosome


SK30
SK33
SK34
SK35
SK36
SK37
SK38
SK40
SK41
SK43
SK45
SK46
Log R Ratio
MAF
Chromosome


SK47
SK48
SK49
Log R Ratio
MAF
Chromosome

B

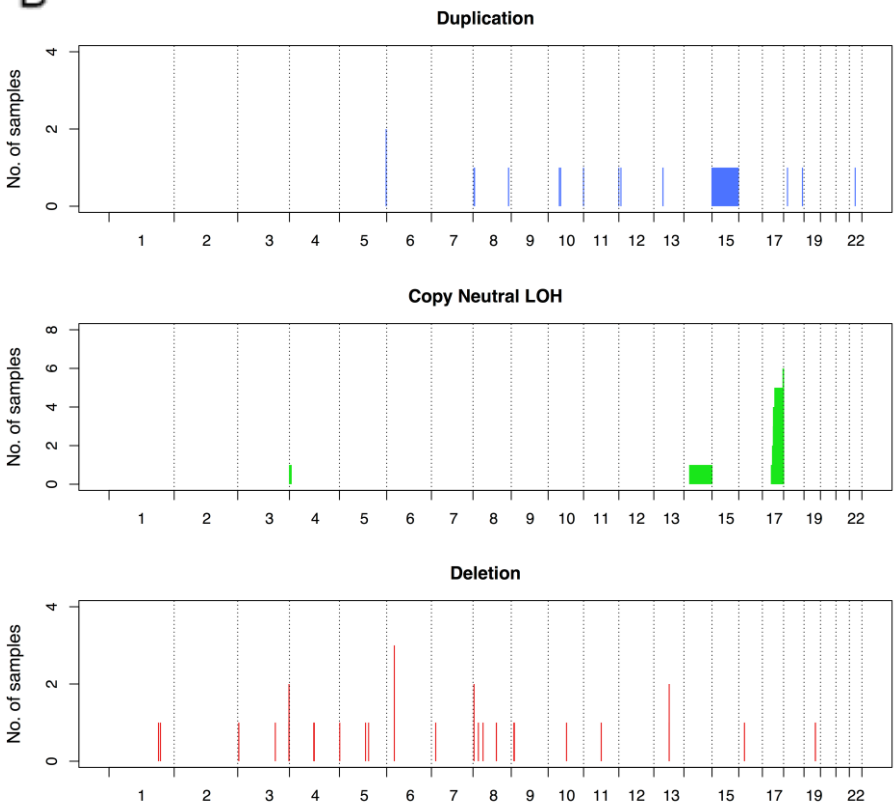

Duplication
Copy Neutral LOH
Deletion
No. of samples
1 2 3 4 5 6 7 8 9 10 11 12 13 15 17 19 22

Figure 6. CNV/LOH called from SNP array of 27 acanthotic sample replication dataset. A. The raw CNV/LOH plot called by minor allele frequency and log r ratio. **B.** Sum of all CNV/LOH called from WES/SNP array of acanthotic SK.

Table 7. CNV/LOH list found in acanthotic samples either WES or SNP array data. Population frequency was obtained from DGV database.

SK	Histology	CNV	Type	DGV	Freq	Gene
SK03a	A	17:41858026-81195210	CN LOH	X	-	
SK03b	A	17:42225547-81195210	CN LOH	X	-	
SK08	A	17:42273501-81195210	CN LOH	X	-	
SK09	A	17:38609187-81195210	CN LOH	X	-	
SK10	A	17:47468020-81195210	CN LOH	X	-	
SK20	A	4:1-6645597	CN LOH	X	-	*FGFR3*
SK21	A	17:34277206-81195210	CN LOH	X	-	
SK22	A	17:78811244-81195210	CN LOH	X	-	
SK48	A	14:21796784-107349540	CN LOH	X	-	

SK09	A	13:58708500-58765648	Deletion	+	-	
SK10	A	19:43390362-43519442	Deletion	+	6.38%	
SK10	A	9:10061106-10132437	Deletion	+	-	
SK10	A	9:11645495-11988769	Deletion	+	0.14%	
SK11	A	3:196130005-196554952	Deletion	-	-	3q29 microdeletion syndrome
SK11	A	4:94986832-95049123	Deletion	-	-	
SK11	A	5:1292983-1344458	Deletion	-	-	*CLPTM1L*(Exon2)-*TERT*(Exon2)
SK11	A	7:16348830-16401190	Deletion	+	0.32%	
SK12	A	4:92906792-93044488	Deletion	+	0.10%	
SK20	A	6:29094696-29161435	Deletion	+	0.77%	
SK20	A	8:20474703-20566971	Deletion	-	-	
SK21	A	6:29094696-29161435	Deletion	+	0.77%	

SK21	A	8:3918118-4032622	Deletion	+	0.03%	
SK22	A	6:29094696-29161435	Deletion	+	0.77%	
SK22	A	8:3918118-4032622	Deletion	+	0.03%	
SK24	A	3:143772247-143858564	Deletion	+	-	
SK26	A	10:70189934-70271354	Deletion	-	-	*SLC25A2*(Exon2)-*DNA2*(Exon15)
SK30	A	3:4041589-4245833	Deletion	+	0.09%	
SK36	A	1:196823300-196901753	Deletion	+	3.40%	
SK36	A	5:111636632-111688150	Deletion	+	-	
SK36	A	8:89354621-89628612	Deletion	-	-	
SK37	A	13:58709338-58765648	Deletion	+	-	
SK40	A	11:67501626-67731956	Deletion	+	0.45%	
SK43	A	16:21608472-21839340	Deletion	+	0.17%	

SK43	A	8:38368270-38419633	Deletion	-	-	*C8orf86*
SK45	A	3:195942132-196519209	Deletion	-	-	3q29 microdeletion syndrome
SK47	A	1:189337997-189542729	Deletion	+	0.26%	
SK47	A	5:99943991-100004949	Deletion	+	-	
SK49	A	19:43328006-43774853	Deletion	+	1.39%	
SK03b	A	5:178730062-178946679	Duplication	+	0.07%	
SK09	A	10:135091802-135534747	Duplication	+	0.19%	
SK22	A	12:8001932-8097433	Duplication	+	3.24%	
SK24	A	11:134869557-134934063	Duplication	+	-	
SK30	A	10:42682133-43043832	Duplication	+	0.10%	
SK34	A	5:178758909-178918122	Duplication	+	0.07%	
SK36	A	18:14747958-14954303	Duplication	+	-	

SK38	A	15:1-100912913	Duplication	-	-	
SK38	A	8:3773842-5950845	Duplication	+	-	
SK40	A	13:34708238-35135179	Duplication	-	-	*LINC00457*
SK40	A	22:22314463-22584809	Duplication	+	4.09%	
SK41	A	10:47543322-47696414	Duplication	+	7.17%	
SK47	A	18:72122611-72345890	Duplication	+	-	
SK47	A	8:135918442-136470247	Duplication	-	-	*LOC286094*

4. Clonal evolution of *ZNF750* revealed by laser microdissection.

From NGS and Sanger sequencing data, it is evident that two genetic event is required and is sufficient to form SK. However, there is no evidence to find out which one of the two mutation hits first. The clonal analysis based on allelic frequency of *FGFR3* and *ZNF750* driver genes shows that almost all cells in a single SK share same mutations. Since it is unlikely that both events take place at a single time point, the precursor with single genetic event is not expanding fast enough to form hyperplastic lesion. The same phenomenon is observed in previous reports, which shows that apparently normal eyelid skin harbors single *PIK3CA* or *FGFR3* mutations (36). Also, there was a report that different order of mutation could result in different histologic type of breast cancer (58), even different disease in myeloproliferative neoplasms (59). Thus I thought that the sequence of genetic events that give rise to SK could provide valuable information for the genetic and molecular mechanism of SK development.

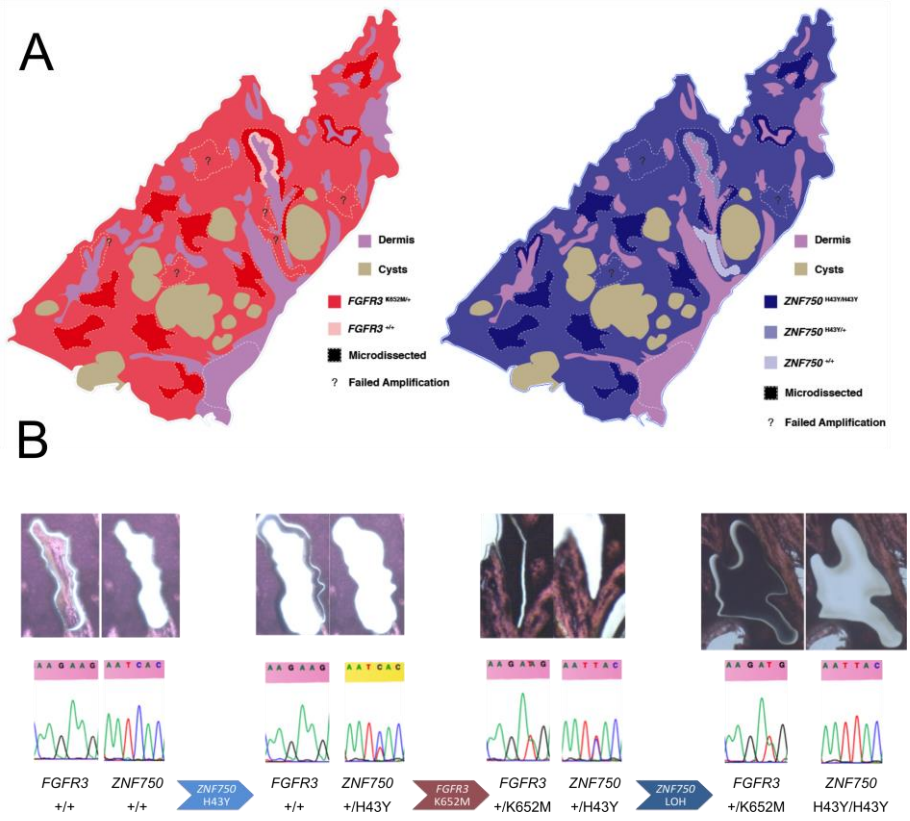
SK tissues proven to have *FGFR3*/*ZNF750* mutations were selected and laser-microdissected. Selective Sanger sequencing was done for specific location of the mutation. The DNA extracted from tissues showed remarkable correlation with expected DNA proportion. Laser microdissection spot was carefully selected to represent histologic layers of epidermis. Almost all

dermal microdissection showed normal genotype in *FGFR3* and *ZNF750*. Some samples showed negligible small peak, which might have come from accident incorporation of nearby SK cells. Microdissection and Sanger sequencing with pure keratinocyte always showed 100% *ZNF750* knock out, in case of *ZNF750* KO samples, and reasonable proportions of *FGFR3*.

Repeated microdissection revealed that SK was consisted by keratinocytes with genetically different population. One sample showed 1:1 *ZNF750* Sanger sequencing read, with wild type *FGFR3*. Subsequent search nearby revealed a cell group that shows 1:1 proportion of wildtype and mutated reads in both *FGFR3* and *ZNF750* genes (Figure 7A). The clonal evolution sequence starts with a single LoF mutation in a *ZNF750* allele, then *FGFR3* activation mutation, and lastly LOH in chromosome 17 for this sample (Figure 7B). The small specific location, which the founding population resides, supports hypothesis that single activating mutation of *FGFR3* is not strong enough to drive SK. No such population was found in other place in the same section, or other section from the same sample.

Clonal evolution of SK could also searched by WES DNA proportion. However, it has a few confounding factors. First, our histologic confirmation step using half of the sample can cause loss of progenitor cells by chance. Second, WES coverage is low in order to effectively assess progenitor with low portion. It is the region why I used laser microdissection on finding clonal evolution sequence. The result, however, seems to correlate well with the

observation that *FGFR3* activation occurs prior to *ZNF750* loss of function in both allele (Figure 7C).


A
Dermis
Cysts
FGFR3 K652M/+
FGFR3 +/+
Microdissected
? Failed Amplification
Dermis
Cysts
ZNF750 H43Y/H43Y
ZNF750 H43Y/+
ZNF750 +/+
Microdissected
? Failed Amplification
B
FGFR3 +/+
ZNF750 +/+
ZNF750 H43Y
FGFR3 +/+
ZNF750 +/H43Y
FGFR3 K652M
FGFR3 +/K652M
ZNF750 +/H43Y
ZNF750 LOH
FGFR3 +/K652M
ZNF750 H43Y/H43Y

C

Sample	Normal	F	FZ	FZZ
SK03a	2.60%	4.71%	23.39%	69.31%

Sample	Normal	Z	ZF	ZFZ
SK03b	4.92%	8.76%	8.32%	78.00%

Sample	Normal	Z	ZF	ZFZ
SK09	11.90%	4.46%	15.36%	68.28%

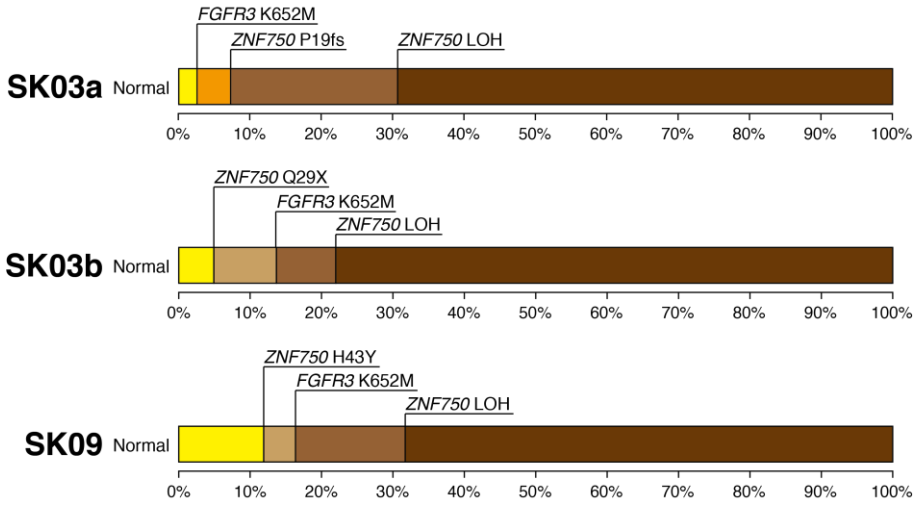


Figure 7. Laser microdissection and clonal evolution map of *FGFR3* and *ZNF750*. A. Schematic drawing of cryosected tissue of SK10. Total 19 dermal and epidermal samples were microdissected and PCR-amplified. A different clone of *FGFR3* and *ZNF750* proportion was found that represents specific steps of *FGFR3-ZNF750* mutation in SK evolution pathway. **B.** Laser microdissection tissue picture and its

Sanger sequencing counterpart. **C.** Result of clonal evolution searched by WES. All three sample *FGFR3* activation followed by *ZNF750* KO.

Discussion

WES of 5 acanthotic SKs and targeted Sanger sequencing of 30 more samples confirmed known driver mutation, *FGFR3*, and another tumor suppressor gene, *ZNF750*, as potential genetic driver in acanthotic subtype of SK.

FGFR3 is the most frequently activated oncogene in acanthotic SK. All 5 of the WES and 70% of the replication samples harbored it. Still, there are acanthotic SK with no *FGFR3* mutation found, and their genetic cause is uncertain.

ZNF750 is a keratinocyte differentiating factor and expressed in epithelial cells of all skin and mucosa (mouth, vaginal, cervix, esophagus) (35). *ZNF750* mutation was previously reported in esophageal squamous cell carcinoma, in the same fashion as SK - LoF associated with LOH, eradication of wild type transcript (54). Just three driver genes, *TP53*, *ZNF750*, and *ATM* can drive malignant esophageal cancer, its potential as tumor driver is well known (17).

Single *FGFR3* mutation does not provide enough proliferative signal to change normal skin to SK. Both normal skin sampling (36) and cell line experiment suggests *FGFR3* activating mutation cause at best mild hyperplasia (30). This hypothesis fits well with *ZNF750* as a secondary hit in acanthotic SK.

Other genetic finding adds more value to this finding. Low number of somatic mutation in some sample, as low as 10, was sufficient to drive SK. Large scale genetic change in chr17q was associated with *ZNF750* also happened, but finding of mere 2068bps deletion in *ZNF750* leads to a conclusion that all 17q LOH points to eradication of wild type *ZNF750* transcript. Thus it can be concluded that only 2 hit, *FGFR3* and loss of *ZNF750*, is sufficient to drive acanthotic hyperplasia in human skin.

Sequence of genetic change in cancer is important because it could affect whole phenotype of the disease and provide valuable information on disease pathogenesis. Such an example is shown in hematopoietic malignancy, that different sequence of two mutations in *JAK2* and *TERT1* resulted in different disease phenotype (59). Another example is LOH event in germline *BRCA1* mutated patients. If loss of *BRCA1* occurs first, cell undergoes apoptosis. Only in context of preceding growing/immortalization could loss of *BRCA1* have effect on carcinogenesis (58). Especially in this case, one frameshift mutation in a copy of *ZNF750* is known to cause autosomal dominant seborrheic dermatitis (53). Difference in phenotype of single and double knockout of *ZNF750* implicates *FGFR3* as a preceding factor. Laser microdissection of acanthotic SK revealed the same relationship in *ZNF750* and *FGFR3*.

Moreover, negative result in complete knockout of CRISPR in HaCaT cell line also backs up the sequential tumorigenic process (data not shown). I

bought pre-designed *ZNF750* CRISPR sgRNA targeting C2H2 conserved region (53). Four times of knock out experiment all failed showing frame conserved deletion outside of C2H2 motif. This negative result supports loss of *ZNF750* in non-carcinoma cell line cause cell cycle arrest. Further knock out of CRISPR experiment with *FGFR3* activating mutation induced HaCaT cell line will add another strong evidence in *FGFR3-ZNF750* tumor evolution pathway shown by laser microdissection (CRISPR performed and permitted by Jiho Park and Hae Yong Yoo, Samsung Advanced Institute for Health Sciences & Technology.).

Some question about secondary genetic change in acanthotic SK remains elusive. However, there are some clues. Rare CNV found by SNP array gives another candidate genes of secondary hits in acanthotic SKs. Other candidate gene is *PAK2*, which is associated in 3q29 deletion of 2 samples SK06 and SK45. *PAK2* gene participate in mTOR pathway (60), downstream molecule of *CDC42*/*RAC1* (61) which interacts with *PIK3CA* (62) and *HRAS* (63) downstream signaling. *HRAS* and *PIK3CA* that are both found to be activated in SK (64). *PAK2* expression level is known to affect clinical outcome of endometrial cancer (65), melanoma (66), and gastric cancer (67). *TERT1* promoter region deletion is also another candidate of secondary hit, which is repeatedly mutated in SK (64). All these candidates could be confirmed by WES or targeted sequencing of specific region.

In this chapter, I have discussed about the development of acanthotic SK. I successfully re-discovered previously known *FGFR3* mutation, and found new causal genetic change, *ZNF750* knock out, of SK through WES. The mutational sequence of two genetic changes mimic tumorigenic sequence of other malignancy. Further cellular *ZNF750* knock out study with CRISPR under context of *FGFR3* activating mutation will better elucidate the importance of tumor mutation sequence.

Chapter 2

Genetics of hyperkeratotic SK

Results

1. WES of SK reveals causal gene: *FGFR3* and *PIK3CA* & other candidates

Through in-house somatic variant calling pipeline, I identified 585 somatic mutations from 3 hyperkeratotic samples (Figure 3A). The mutational burden for each hyperkeratotic sample seems to be higher than acanthotic samples (195 vs. 58.6 per sample, p-value=5.652 * 10^{-3}), but other factors such as degree of exposure to sunlight (p-value=6.862 * 10^{-3}) and age might contribute to the factor randomly.

Mutations recurrently found in different samples include *FGFR3*, *PIK3CA*, *NOTCH1*, *DNAH5*, *FPR3*, etc. *FGFR3* and *PIK3CA* passed Bonferroni correction in protein alteration binomial test. *FGFR3* and *PIK3CA* are the only genes that passed Bonferroni correction (Table 8).

FGFR3 variant was found in all 3 samples. 3 different mutations, A393E, R248C, S311F were found in *FGFR3*. A393E and R248C were also COSMIC variants. However, S311F is a new variant without known activating potential (Figure 3C, 8A, Table 9).

PIK3CA variants included E542K, I841V, and H1047R mutations. Variants E542K and H1047R are known as on of the most important cancer-causing hotspot in *PIK3CA* gene, and they were also reported to be involved in the development of SK and benign epidermal nevi (8, 32). I841V mutation

is not conserved well in other vertebrates, not found in COSMIC, substituted to same branched-chain amino acid group. Its effect on *PIK3CA* activity is doubtable (Figure 3C, 8B, Table 9).

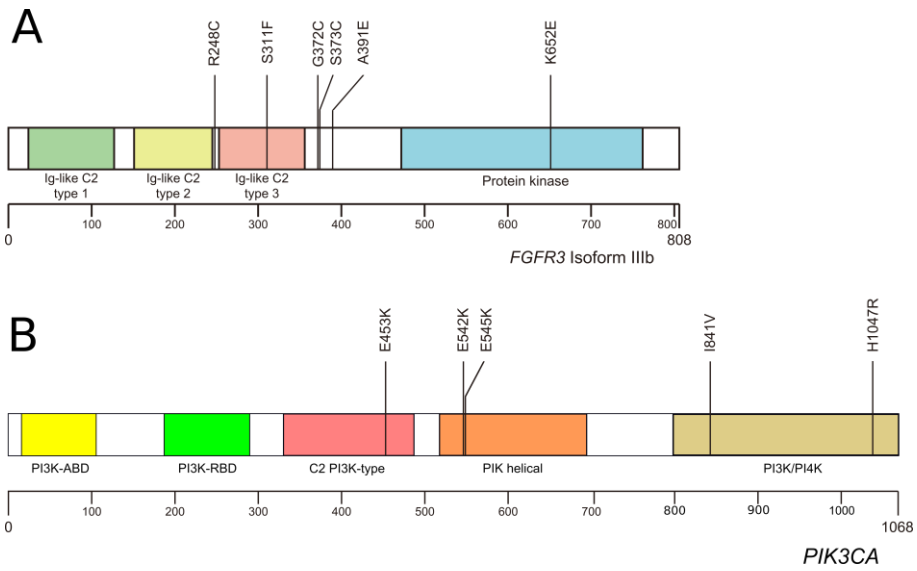


Figure 8. Cancer drivers found in WES of hyperkeratotic samples.

A. *FGFR3* gene structure and somatic mutation found in WES and Sanger sequencing of acanthotic SKs. **B.** *PIK3CA* mutations found in WES and Sanger sequencing of acanthotic SKs.

Table 8. List of recurrently mutated genes in hyperkeratotic SK.

Gene	Total # of Variants	Conserved[6]	Silent	Protein Altering	Damaging[7]	Binomial-total	Binomial-altering	Binomial-conserved	Binomial-damaging
FGFR3	3	3	0	3	0	2.873.E-05‡	9.842.E-06‡	4.111.E-07‡	9.970.E-01
PIK3CA	3	2	0	3	0	1.632.E-04	5.671.E-05‡	2.951.E-04	9.946.E-01
GDF2	2	0	0	2	0	4.399.E-04	2.149.E-04	9.928.E-01	9.984.E-01
ZNF554	2	0	0	2	0	6.866.E-04	3.362.E-04	9.910.E-01	9.980.E-01
DPY19L2	2	1	0	2	0	1.342.E-03	6.603.E-04	1.257.E-02	9.972.E-01
DGKI	2	2	0	2	0	2.593.E-03	1.284.E-03	1.572.E-04	9.960.E-01

[6] Amino acid sequence difference equal or less than 2 in 100-species alignment.

[7] Framshift, splicing, or nonsense variants.

‡ Passed Bonferroni test.

TRPA1	2	0	0	2	0	2.852.E-03	1.414.E-03	9.814.E-01	9.958.E-01
PLCB1	2	1	0	2	0	3.345.E-03	1.662.E-03	2.002.E-02	9.955.E-01
FMN2	2	0	0	2	0	6.474.E-03	3.251.E-03	9.715.E-01	9.936.E-01
RP1	2	1	0	2	0	9.845.E-03	4.990.E-03	3.493.E-02	9.920.E-01
NOTCH1	2	1	0	2	1	1.344.E-02	6.872.E-03	4.112.E-02	9.415.E-03
MED6	2	1	1	1	0	1.465.E-04	1.187.E-02	4.116.E-03	9.991.E-01
FPR3	3	0	2	1	0	2.472.E-06‡	1.695.E-02	9.941.E-01	9.987.E-01
DNAH17	2	1	0	2	0	3.586.E-02	1.910.E-02	6.955.E-02	9.835.E-01
DNAH5	3	1	1	2	0	4.120.E-03	2.035.E-02	7.188.E-02	9.829.E-01
CEL	2	0	1	1	0	1.335.E-03	3.558.E-02	9.874.E-01	9.972.E-01
CCDC168	2	0	0	2	1	7.515.E-02	4.232.E-02	8.878.E-01	2.566.E-02
CNTN4	2	0	1	1	0	2.413.E-03	4.766.E-02	9.829.E-01	9.962.E-01
DNAH9	2	0	1	1	0	3.619.E-02	1.760.E-01	9.274.E-01	9.835.E-01

TTN	2	1	0	2	0	2.635.E-01	2.604.E-01	3.203.E-01	8.831.E-01
GPLD1	2	0	2	0	0	1.639.E-03	9.598.E-01	9.860.E-01	9.969.E-01

Table 9. WES result of oncogenic candidates of Hyperkeratotic SK.

Sample ID	Gene	Chromosome: position (hg19)	Base change	Somatic mutation	Impact on protein	Amino acid change	Amino acid location / protein length	COSMIC gene/ position	# of species different from human/ # of species with ortholog	Coverage depth (alternative allele/total)	
										Tumor	Normal
SK_01	*FGFR3*	Chr4:1803564	C>T	AA>AB	Missense	R248C	248/808	True/True	0/0	16/46 (34.8%)	0/40 (0%)
SK_02	*FGFR3*	Chr4:1806153	C>A	AA>AB	Missense	A393E	391/808	True/True	45/0	28/74 (37.8%)	1/58 (1.7%)
SK_06	*FGFR3*	Chr4:1804642	C>T	AA>AB	Missense	S311F	311/808	True/na	2/0	27/108 (25.0%)	0/64 (0%)
SK_01	*PIK3CA*	Chr3:178936082	G>A	AA>AB	Missense	E542K	542/1068	True/True	1/0	21/89 (23.6%)	0/85 (0%)

SK_02	*PIK3CA*	Chr3:178952085	A>G	AA>AB	Missense	H1047R	1047/1068	True/True	1/0	207/556 (37.2%)	0/258 (0%)
SK_06	*PIK3CA*	Chr3:178947085	A>G	AA>AB	Missense	I1841V	841/1068	True/na	6/0	45/159 (28.3%)	0/63 (0%)

2. *FGFR3* and *PIK3CA* of hyperkeratotic SK in extended data set

FGFR3 and *PIK3CA* seem to be the major driver of hyperkeratotic SK. I Sanger sequenced two genes in 13 more hyperkeratotic SK and its normal salivary counterparts (Table 5).

6 of 13 samples were found to have *FGFR3* mutation, which are R248C (2 samples), G372C, Y375C, and K652E. 3 of 13 samples harbored *PIK3CA* mutation, which are H1047R/L (each 1 sample) and E542K. With WES result, hyperkeratotic SKs harbored 56.3% (9 of16) *FGFR3*, and 37.5% (6 of 16, including I841V) *PIK3CA* mutations. Both genes, with extended sequencing data, passed genome-wide significance level in binomial test (Table 10).

All *PIK3CA* mutations occurred together with *FGFR3*. *PIK3CA* E545K and H1047R mutations were also found in acanthotic samples, but in much lower frequency (2 of 35 samples, 5.7%).

Table 10. P-value change of *FGFR3* and *PIK3CA* after Sanger sequencing 13 replication samples. Both genes passed genome-side significance level of 10^{-6}.

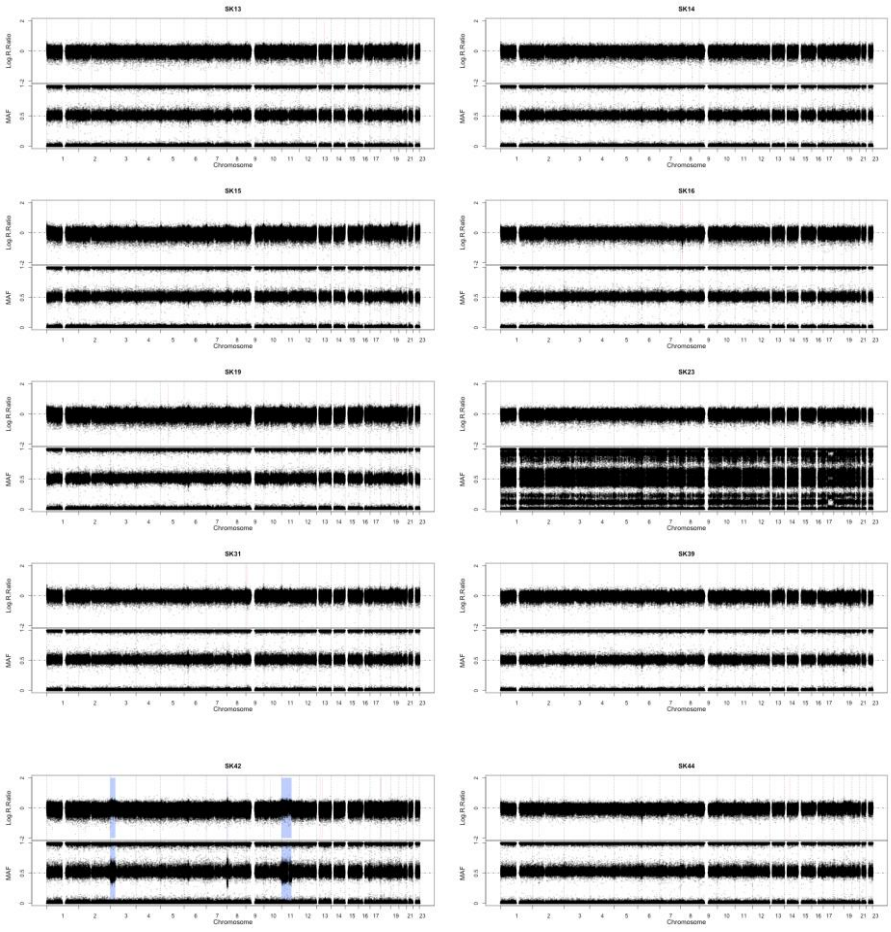
Genes	Binom-total	Binom-alter	Binom-con
FGFR3	1.28466E-09	7.73153E-11	1.73975E-14
PIK3CA	2.17675E-05	2.91653E-06	2.83817E-07

3. CNV confirmed by SNP array

Copy neutral (CN) LOH and CNV events were called from coverage information of WES. SK06 sample had CN LOH in 9q, deletion at 10q regions (Figure 4, Table 11).

In replication data set, I searched CNV of 10 hyperkeratotic SK samples with SNP array (Figure 9A). There was no recurrently found CNVs (Figure 9B, Table 11). Found CNVs include duplication in chr11p, 3p (SK42). Microdeletion including *RB1* gene at chr13 (SK42), microduplication including *MSH2* and *MSH6* gene at chr2 (SK44) are also notable.

A


SK13
SK14
SK15
SK16
SK19
SK23
SK31
SK39
SK42
SK44
Log R Ratio
MAF
Chromosome

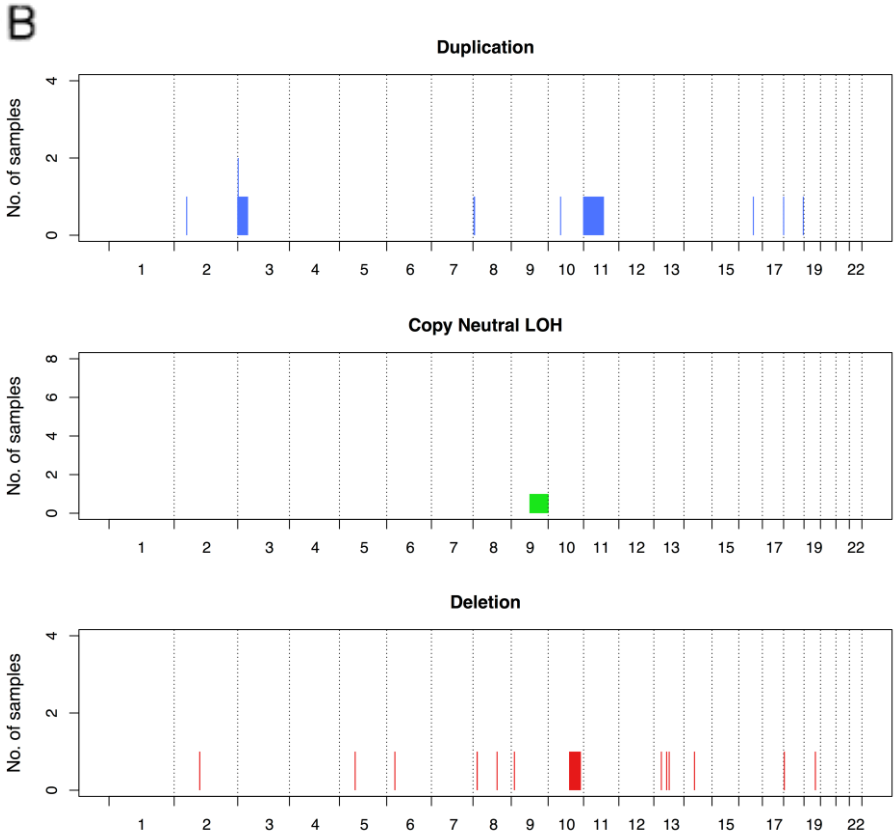


Figure 9. CNV/LOH called from SNP array of 9 hyperkeratotic sample replication dataset. A. The raw CNV/LOH plot called by minor allele frequency and log r ratio. **B.** Sum of all CNV/LOH called from WES/SNP array of hyperkeratotic SK.

Table 11. CNV/LOH list found in hyperkeratotic samples either WES or SNP array data. Population frequency was obtained from DGV database.

SK	Histology	CNV	Type	DGV	Freq	Gene
SK06	H	9:71555522-141213431	CN LOH	X	-	
SK06	H	10:82012954-123353315	Deletion	-	-	
SK13	H	13:58708500-58765648	Deletion	+	-	
SK13	H	2:97764450-98024837	Deletion	+	0.53%	
SK14	H	8:92128840-92181214	Deletion	+	0.16%	
SK16	H	8:15692532-16212165	Deletion	-	-	*MSR1*
SK19	H	19:43390362-43559330	Deletion	+	4.87%	
SK19	H	5:60169151-60238590	Deletion	+	0.02%	
SK31	H	9:12013711-12108916	Deletion	+	1.05%	
SK42	H	13:28814370-28868293	Deletion	+	-	

SK42	H	13:48902121-48963784	Deletion	-	-	*RB1*
SK42	H	18:2642376-2804129	Deletion	+	-	
SK44	H	14:40296775-40364644	Deletion	+	-	
SK44	H	6:31360255-31453640	Deletion	+	6.24%	
SK16	H	3:1752478-1840129	Duplication	+	0.05%	
SK19	H	16:55781785-55856717	Duplication	+	-	
SK39	H	18:75240052-75395691	Duplication	+	-	
SK42	H	10:47605311-47703869	Duplication	+	7.30%	
SK42	H	11:1-76047793	Duplication	-	-	
SK42	H	18:87505-196829	Duplication	+	0.05%	
SK42	H	3:1-37705729	Duplication	-	-	
SK42	H	8:3686944-5926872	Duplication	+	0.35%	
SK44	H	2:47890331-48013099	Duplication	-	-	*MSH2*(Exon18~19),

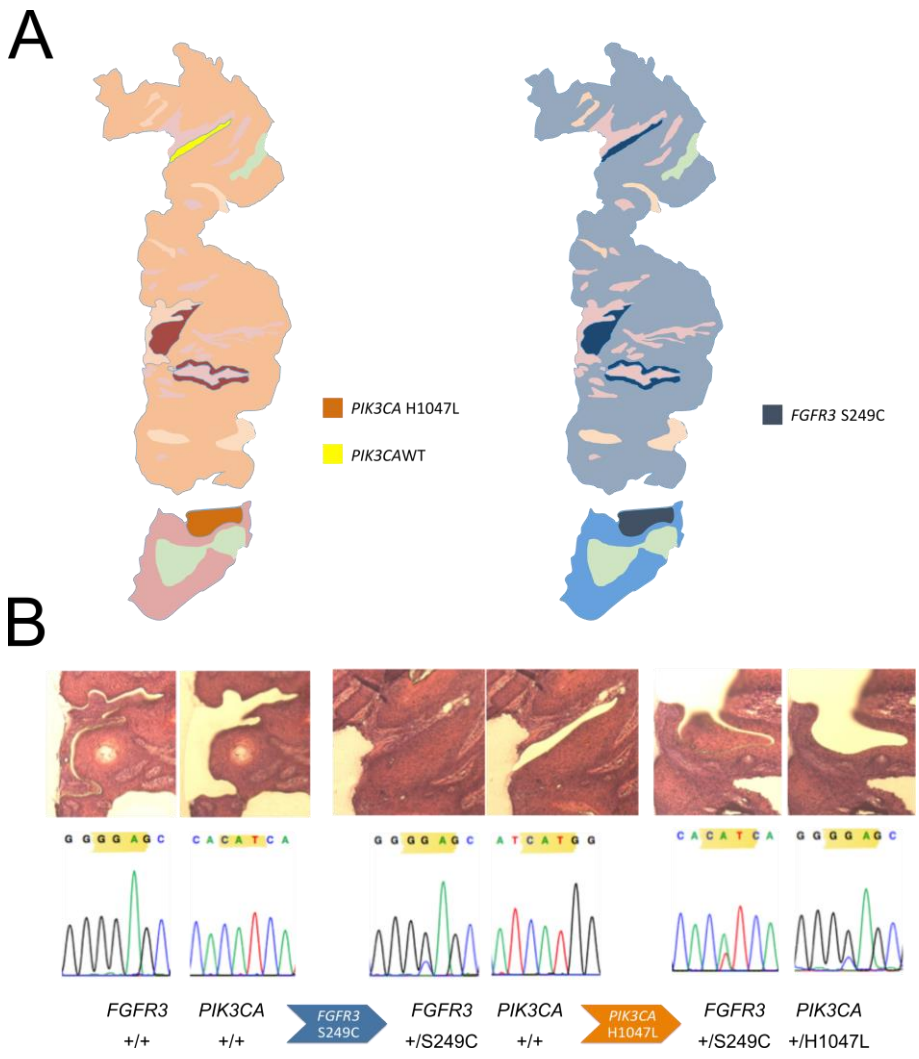
MSH6(Exon1~2)

4. Clonal evolution of *FGFR3* and *PIK3CA* revealed by WES and laser microdissection

Laser microdissection of a sample with *PIK3CA* revealed another founding clone. *PIK3CA* and *FGFR3* mutations were found in 1:1 proportion in every other dissection. A single section shows different result, without *PIK3CA* mutation but *FGFR3* mutation (Figure 10A). From this sample, it is implicated that *FGFR3* activation comes first, than activation of *PIK3CA* (Figure 10B). Clonal evolution suggested by WES also follows *FGFR3* mutated first in SK01 and SK02 samples. In SK06 sample, *PIK3CA* occurs faster than *FGFR3*. This might be due to its less potential driving force, or due to random sequencing error (Figure 10C). *NOTCH1* seems to play part in the evolution of hyperkeratotic SK.

In acanthotic sample, the percentage of cells with *ZNF750* and *FGFR3* mutation found in WES was same in all three samples; so only laser microdissection could reveal its mutational sequence. In hyperkeratotic SK, SK01 showed interesting difference in percentage of *FGFR3* and *PIK3CA* mutation proportion. 69.5% of the cells harbored *FGFR3* R248C, while only 47.2% of the cells harbored *PIK3CA* E542K mutation. If these two mutations were the only drivers in SK01, it would contradict to previous observation that single *FGFR3* activation is not enough to drive hyperplasic event in vivo(34, 36). In SK01, one splicing variant and one conserved-site missense

mutation of *NOTCH1* was harbored by 70.1% of the cells, proportion strikingly similar to that of *FGFR3*. *NOTCH1* is frequently mutated in early development of cutaneous SCCs (68).


A
PIK3CA H1047L
PIK3CAWT
FGFR3 S249C
B
FGFR3
+/+
PIK3CA
+/+
FGFR3
S249C
FGFR3
+/S249C
PIK3CA
+/+
PIK3CA
H1047L
FGFR3
+/S249C
PIK3CA
+/H1047L

C

Sample	Normal	N1	N2	NNF	NNFP
SK01	29.71%	0.16%	0.56%	22.56%	47.20%

Sample	Normal	P	PF	PFN	PFNN
SK06	43.40%	6.60%	9.10%	3.40%	37.51%

Sample	Normal	F	FP
SK02	24.40%	1.20%	74.40%

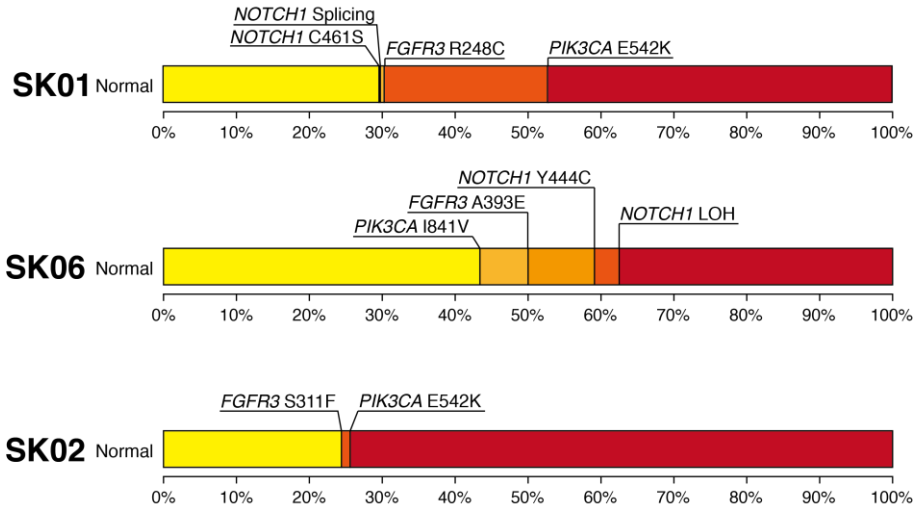


Figure 10. Laser microdissection and clonal evolution map of *FGFR3* and *PIK3CA*. A. Schematic drawing of cryosected SK50 sample. Total 11 dermal and epidermal samples were microdissected and PCR-amplified. A different clone of *FGFR3* and *PIK3CA* proportion was found that represents specific steps of *FGFR3-PIK3CA* sequential mutation in SK evolution pathway. **B.** Laser microdissection tissue picture and its Sanger sequencing counterpart. **C.** Result of

clonal evolution searched by WES. Two samples have *FGFR3* activation, followed by *PIK3CA* activation mutation.

5. Histologic subtypes of SK is determined by genetic profile

There was no study that found the basis for SK histologic difference. While investigating mutational distribution across histologic subtype, I observed two histologic subtypes of SK harbors different mutational profile. In WES data, every *ZNF750* knockout mutations were found in acanthotic SK, and *PIK3CA* mutations were found in hyperkeratotic subtypes. This tendency remained throughout follow-up samples, 6 *ZNF750* mutations were all found in acanthotic SK, 5 of 7 *PIK3CA* mutations were found in hyperkeratotic SK. Two *PIK3CA* E453K and 1047L mutations were found in an acanthotic samples, but overall trend not affected. This distribution reached statistical significance (p-value = $4.777*10^{-3}$).

6. Exposure to UV and increasing age as risk factors of SK

The number of somatic mutations widely varied from sample to sample in WES: 10 mutations for the least, and 279 the most. The samples had 60~70% ratio of UV signal mutation, except for SK03b sample acquired from belly which had 30% (Figure 11A, Table 12). This might be due to low exposure to sunlight, but it may also be the result of low number of somatic mutation it harbors (binomial test for UV signature cutoff 60%, p = 0.1011). Number of UV induced somatic mutations is higher in samples from exposed body area (p-value = $2.796*10^{-3}$), and higher with longer exposure time, which is increasing age (R^2 = 0.4622, p-value = $3.81*10^{-2}$). However, the result for UV exposure area and somatic mutation number should be take carefully and backed up with more samples. Many other factors could affect the result: sunscreen usage, personal and social interaction frequency, indoor/outdoor job, etc. The age distribution between exposed and non-exposed group was not uniform as well (p-value = 0.0469), by chance due to low number of samples sequenced. WES of more samples, with their social and personal characteristics would provide better association without any bias.

Another trend was found in UV characteristics. As subject's age increased, total number of non-UV induced somatic mutation in SK samples increased (Figure 11B, R2 = 0.5829, p-value = $1.67*10^{-2}$), showing additional

contribution of increasing age in SK on top of longer UV exposure.

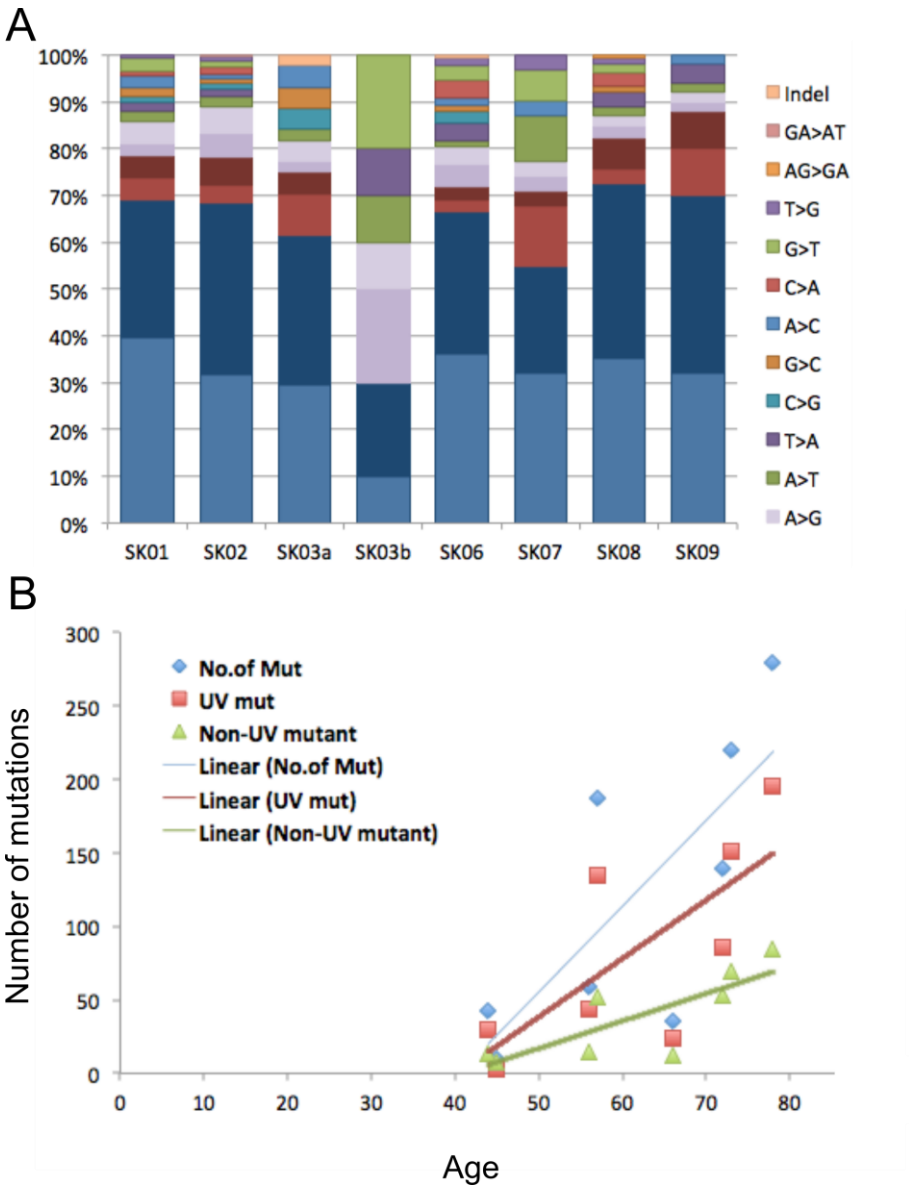


Figure 11. UV-associated base change and somatic mutation trend of SK.

A. The base change pattern of SK somatic mutation follows typical UV

associated cutaneous malignancies. **B.** Somatic mutation number is positively correlated with age increase. Both UV and non-UV associated base change increase as subject age increase, implying independent effect of age.

Sample	C>T % at YC:RG>YT:RA	CC>TT proportion	Total UV signature
SK_01	(117/206) 56.8%	(10/206) 9.7%	(127/206) 66.5%
SK_02	(159/270) 58.9%	(26/270) 9.6%	(185/270) 68.5%
SK_03a	(24/43) 55.8%	(6/43) 14.0%	(30/43) 69.8%

Table 12. Proportion of UV signature mutations from 8 WES samples.

SK_03b	(3/10) 30%	(0/10) 0%	(3/10) 30%
SK_06	(71/139) 51.1%	(14/139) 10.07%	(85/139) 61.2%
SK_07	(14/36) 38.88%	(10/36) 27.77%	(24/36) 66.7%
SK_08	(97/187) 51.87%	(38/187) 20.32%	(135/187) 71.4%
SK_09	(26/59) 44.06%	(18/59) 30.5%	(44/59) 74.6%

7. Recurrent mutations in both histologies

In order to find more oncogenic driver candidate in SK, searching for recurrent mutation regardless of histology seemed feasible. *PIK3CA* mutation found in both histologies could raise possibility that same genetic pathway could be shared as well.

Although there were no inter-histologic mutations that passed bonferroni correction, some interesting examples were found. *FRS3* mutation, found in SK02 and SK08, participate in downstream signal pathway of fibroblast grow factor. *FRS3* forms complex with *FRS2*, transduces FGFR receptor family signal to both MAPK and PI3K pathway (69). Mutation in this gene might affect *FGFR3* and *PIK3CA* signal cascade.

ZNF534 gene has another interesting feature. It is mutated twice in WES samples, but both mutations at the same location, E483K. The role of *ZNF534* and effect on SK development is largely unknown. The same variation is found in ExAC database, but in very low frequency. Searching for *ZNF534* mutation with bigger samples in WES might reveal its unrecognised role in SK tumorigenesis.

Table 13. List of inter-histologic recurrent genes.

Gene	Total	Conserved	Silent	Protein alter	Damaging	Binom-total	Binom-alter	Binom-con	Binom-damaging
PLCB1	3	1	0	3	0	3.083.E-04	1.081.E-04	1.149.E-01	9.726.E-01
COL5A2	3	1	1	2	0	5.606.E-04	5.412.E-03	1.114.E-01	9.736.E-01
ALDOB	2	0	1	1	1	7.077.E-04	2.599.E-02	9.454.E-02	2.176.E-02
PYHIN1	2	0	0	2	0	1.081.E-03	5.313.E-04	7.805.E-01	9.489.E-01
CLEC14A	2	0	0	2	0	1.265.E-03	6.226.E-04	9.584.E-01	9.910.E-01
FRS3	2	0	0	2	0	1.276.E-03	6.276.E-04	9.661.E-01	9.927.E-01
FSHR	2	0	1	1	0	2.315.E-03	4.671.E-02	9.663.E-01	9.928.E-01
ZNF534	2	0	0	2	0	2.348.E-03	1.162.E-03	8.940.E-01	9.765.E-01
DDX31	2	0	1	1	1	3.675.E-03	5.862.E-02	1.264.E-02	2.705.E-03
RAD54B	2	1	1	1	0	4.176.E-03	6.242.E-02	3.383.E-02	9.926.E-01

ZBED4	2	0	0	2	0	6.727.E-03	3.383.E-03	9.788.E-01	9.955.E-01
MPHOSPH9	2	0	0	2	0	6.857.E-03	3.450.E-03	9.862.E-01	9.971.E-01
NRCAM	2	1	1	1	0	7.165.E-03	8.125.E-02	1.202.E-01	9.712.E-01
RASGRF2	2	1	0	2	0	7.455.E-03	3.757.E-03	2.822.E-02	9.939.E-01
ANK2	3	1	1	2	1	7.960.E-03	3.149.E-02	2.410.E-05	1.474.E-03
ZNF521	2	1	1	1	0	8.309.E-03	8.732.E-02	8.358.E-02	9.808.E-01
CNTNAP2	2	1	0	2	0	8.546.E-03	4.319.E-03	4.087.E-02	9.910.E-01
CPS1	2	2	0	2	0	1.074.E-02	5.460.E-03	7.107.E-03	9.734.E-01
PCF11	2	0	0	2	0	1.139.E-02	5.800.E-03	9.788.E-01	9.955.E-01
HYDIN	3	0	1	2	1	1.526.E-02	4.843.E-02	4.214.E-02	9.243.E-03
PTPRB	2	2	0	2	0	1.651.E-02	8.500.E-03	1.761.E-04	9.960.E-01
MYO18B	2	0	1	1	0	2.791.E-02	1.559.E-01	9.636.E-01	9.922.E-01

UNC80	2	0	2	0	0	4.129.E-02	7.889.E-01	8.183.E-01	9.584.E-01
IGFN1	2	0	2	0	0	5.163.E-02	7.619.E-01	9.897.E-01	9.978.E-01
RYR3	2	2	0	2	0	7.880.E-02	4.461.E-02	2.457.E-03	9.847.E-01
AHNAK	2	0	0	2	0	1.035.E-01	6.055.E-02	9.746.E-01	9.946.E-01

Discussion

WES and Sanger sequencing of 16 hyperkeratotic SK confirmed previously reported *FGFR3* and *PIK3CA* mutation.

FGFR3 mutation frequency was lower in hyperkeratotic SK in replication samples compared to acanthotic subtype, about 60%. *FGFR3* mutation profile was more variable than acanthotic subtypes.

PIK3CA is known as 2nd most commonly mutated gene after *FGFR3* in SK. 3 samples of WES and 13 more replication samples successfully confirmed *PIK3CA* as an oncogene of hyperkeratotic SK. 4 of 5 *PIK3CA* mutations were mutated in mutational hotspot found in COSMIC database, 1 mutation potential is questionable due to its low conservation. *PIK3CA* was mutated in 25% of all hyperkeratotic SKs.

Observation of segregation pattern revealed possible genetic-histologic relationship of SK. Similar association is observed in breast cancer. *PTEN*, *TP53*, and *BRCA1* LOH is key three molecular change of breast cancer carcinogenesis. The primary genetic change, *PTEN* or *TP53* loss, dominates histologic and clinical outcome (58). Likewise, WES and Sanger sequencing data infer histologic difference driven by genetic change.

This example gives idea that SK could also share conventional pathways of malignant skin cancers is shown with SK01 and SK06 samples. *NOTCH1* is a tumor suppressor gene, frequently mutated in other epithelial

squamous cell carcinoma such as esophageal SCC, cutaneous SCC, head and neck SCC, bladder SCC (68, 70). *NOTCH1* knock out mouse shows various skin malignancies as well as many dysplastic nodules. Hyperplasic change in cornea of KO mouse histologically resembles that of SK (71). In my sample, SK01 had one *NOTCH1* loss of function mutations and a slicing variant. SK06 had LOH involving *NOTCH1* region. For SK01, *NOTCH1* shows higher prevalence in cancer cell than *PIK3CA*, showing independent role in cellular hyperplasia and its possibility as 3rd tumor driver in hyperkeratotic SK. WES of more samples is needed to reveal its prevalence in hyperkeratotic SK more thoroughly. Also, B-allele difference in a WES sample implicated *NOTCH1* as possible driver mutation.

Laser microdissection data successfully revealed each clonal evolution history in both histologic subtypes. One important observation in all three samples is that mutation in *FGFR3* always precedes other mutations. In *ZNF750* case, it precedes LOH of *ZNF750*. This kind of initiating mutation is found in large-scale genetic study of other malignancies: breast (58), cutaneous SCC (24), and esophageal cancer (17). First mutation acts as a master key to determine a cancer's histologic subtype (58), allows normally apoptotic mutation to occur (72). Along with previous reports that linked *FGFR3-FOXN1* signal as a key to benign nature of SK (29, 42), I also added another layer of understanding that *FGFR3* could be the determinant of cell

fate in SK. Further study with WES and RNASeq of more samples, cell line experiments are required to find the role of FGFR3 in SK.

The presence of *PIK3CA* E542K and H1047R mutation in two of acanthotic samples shows possibility that some other factors could also contribute to histology of SK. WES and RNASeq could reveal possible mechanism of such genetic-histologic mismatch. One possibility is sequence of genetic event. In breast cancer, major factor of histologic differentiation was first genetic event. Even if same genetic change that determined histologic fate in first step occurred, it did not change the outcome (58).

In my data set, no sample with both *ZNF750* knockout and *PIK3CA* activation was observed. This segregation might come from probability due to low sample number (binomial p-value=0.630). In esophageal carcinoma, sample size of 144 showed 2 samples that harbored both *ZNF750* and *PIK3CA* mutations (p-value = 1) (54). Assuming same prevalence of *ZNF750* and *PIK3CA*, at least 175 SK samples are needed to show statistic significance of mutual exclusive occurrence of *PIK3CA* and *ZNF750* in SK.

These information successfully revealed driver genes of hyperkeratotic SK. Activation of *FGFR3* and *PIK3CA* through two genetic events is sufficient for its development. The incidence of genetic change depends on degree of UV exposure and patient's age.

Conclusion

Through the study, genetic landscape and evolution pathway of SK development were discovered. WES of 8 pilot SK samples could validate previously known drivers of SK; *FGFR3* and *PIK3CA*. New gene *ZNF750* was also discovered as profound driver of SK. Follow-up sanger sequencing study with 44 additional samples showed their exact proportion in SK. Altogether, 70% of the SK could be explained by *FGFR3*, *PIK3CA* and *ZNF750* alone. Other genes such as *NOTCH1*, 3q29 microdeletion, *RB1*, *FRS2*, and *ZNF534* were potential drivers of SK, but further WES with larger samples should be done to elaborate their role (Figure 12A). Histologic differentiation of SK was also discovered through WES and follow-up study. *FGFR3* activation and *ZNF750* KO mutation was highly correlated with acanthotic subtype of SK, while *FGFR3* activation and *PIK3CA* activation with hyperkeratotic subtype. This is another example of genetic-histologic segregation. Sequence of mutation in tumor development in SK was also discovered in detail. Clone proportion analysis was done through WES minor allele coverage data. Random bias by WES itself and clonal loss due to initial histologic confirmation was overcome with replication laser microdissection study. The result sequence of mutation was successfully replicated in laser microdissection study. Of 10 cases of SK examined, 6 throught WES and 2

through laser microdissection, *FGFR3* activation mutation occurs prior to *ZNF750* KO and *PIK3CA* activation mutation. One exception was with a sample with *PIK3CA* mutation prior to *FGFR3*, but with questionable activation potential I841V. This could be just a passenger mutation. Base change analysis showed independent effect of increasing age and UV exposure on SK tumorigenesis, suggesting genetic risk factors of SK (Figure 12B).

The limitations of the study are: first, low number of WES samples. Statistical analysis was made based on assumption that mutation number of 8 SK samples represent actual mutation number of the SK that were analyzed with targeted sanger sequencing. The limitation could be overcome with more samples of SK searched through WES. Second, there was no explanation for tumorigenesis of SK samples without drivers. WES, as well as whole genome sequencing, RNA sequencing could be done to show exomic and possible noncoding driver mutations. Third, low number of coverage in WES could induce random error in proportion of progenitor cells, confounding the result of clonal evolution. Low percentages of SK progenitor cells enhance this error. Targeted deep sequencing would successfully overcome this error and reveal better structure of clonal evolution in SK. Lastly, *in vitro* and *in vivo* replication study of *FGFR3*, *PIK3CA*, *ZNF750* tumorigenesis effect and evolutionary relationship could be done. *ZNF750* protein expression status in *ZNF750* positive acanthotic SK can reveal possible post-transcriptional and

post-translational control. Pathway analysis, and protein-protein interaction network analysis such as RNAseq and ChIPseq can determine possible underlying molecular pathway under SK drivers. Cell proliferation, colony formation, and invasion assays with *FGFR3*, *PIK3CA*, and *ZNF750* KO mutations would better reveal SK senescence and pathway interaction. *FGFR3* activation mouse and *ZNF750* KO mouse can be made and studied for SK model organism.

In conclusion, the study successfully revealed SK genomic landscape, major drivers and their possible genetic interactions. As a benign neoplasm of keratinocyte, the genetic landscape of SK could provide much more information in cancer development and treatment.

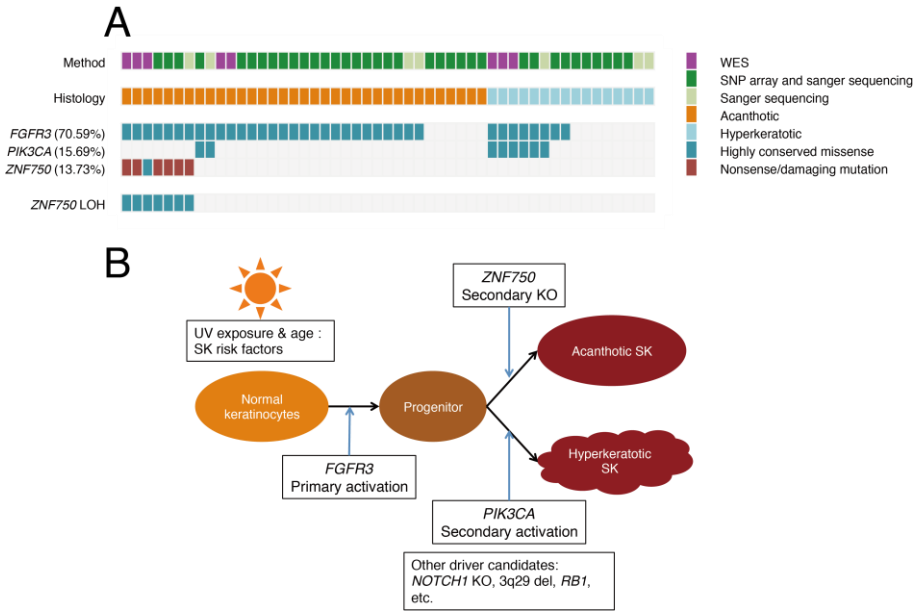


Figure 12. Result summary. A. Through 8 WES and 43 follow-up samples, 3 genes, *FGFR3*, *PIK3CA*, and *ZNF750* were found as genetic drivers of SK. *FGFR3* and *PIK3CA* work as oncogene, while *ZNF750* as a tumor suppressor gene. They totally explain about 70% of whole SK development. **B.** The evolution pathway of SK. UV exposure or increasing age accelerates *FGFR3* activation mutation prior to other mutations, which cause mild increase in proliferation. Than *ZNF750* or *PIK3CA* mutation occurs, driving precursor to either acanthotic or hyperkeratotic SK. Other drivers such as *NOTCH1*, 3q29, *RB1* and their prevalence in SK should be assessed in detail in further studies.

Reference

1. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A. 2009;106(45):19096-101.

2. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455(7216):1061-8.

3. Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015;19(1A):A68-77.

4. Leiserson MD, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet. 2015;47(2):106-14.

5. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474(7353):609-15.

6. International Cancer Genome C, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. Nature. 2010;464(7291):993-8.

7. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546-58.

8. Kato S, Lippman SM, Flaherty KT, Kurzrock R. The Conundrum of Genetic "Drivers" in Benign Conditions. J Natl Cancer Inst. 2016;108(8).

9. Testa U, Pelosi E, Castelli G. Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. Med Sci (Basel). 2018;6(2).

10. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Primers. 2015;1:15065.

11. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. Proc Natl Acad Sci U S A. 2015;112(1):118-23.

12. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-74.

13. Sawyers CL. Chronic myeloid leukemia. N Engl J Med. 1999;340(17):1330-40.

14. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A Big Bang model of human colorectal tumor growth. Nat Genet. 2015;47(3):209-16.

15. Sprouffske K, Pepper JW, Maley CC. Accurate reconstruction of the temporal order of mutations in neoplastic progression. Cancer Prev Res (Phila). 2011;4(7):1135-44.

16. Maley CC, Aktipis A, Graham TA, Sottoriva A, Boddy AM, Janiszewska M, et al. Classifying the evolutionary and ecological features of neoplasms. Nat Rev Cancer. 2017;17(10):605-19.

17. Hao JJ, Lin DC, Dinh HQ, Mayakonda A, Jiang YY, Chang C, et al. Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. Nat Genet. 2016.

18. Gordon R. Skin cancer: an overview of epidemiology and risk factors. Semin Oncol Nurs. 2013;29(3):160-9.

19. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-8.

20. Narayanan DL, Saladi RN, Fox JL. Ultraviolet radiation and skin cancer. Int J Dermatol. 2010;49(9):978-86.

21. Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. Mutat Res. 2005;571(1-2):19-31.

22. Jayaraman SS, Rayhan DJ, Hazany S, Kolodney MS. Mutational landscape of basal cell carcinomas by whole-exome sequencing. J Invest Dermatol. 2014;134(1):213-20.

23. Shain AH, Garrido M, Botton T, Talevich E, Yeh I, Sanborn JZ, et al. Exome sequencing of desmoplastic melanoma identifies recurrent *NFKBIE* promoter mutations and diverse activating mutations in the MAPK pathway. Nat Genet. 2015;47(10):1194-9.

24. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. Clin Cancer Res. 2014;20(24):6582-92.

25. Hafner C, Vogt T. Seborrheic keratosis. Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG. 2008;6(8):664-77.

26. Arbiser JL, Bonner MY. Seborrheic Keratoses: The Rodney Dangerfield of Skin lesions, and Why They Should Get Our Respect. Journal of Investigative Dermatology. 2016;136(3):564-6.

27. Hafner C, van Oers JM, Hartmann A, Landthaler M, Stoehr R, Blaszyk H, et al. High frequency of FGFR3 mutations in adenoid seborrheic keratoses. J Invest Dermatol. 2006;126(11):2404-7.

28. Kwon OS, Hwang EJ, Bae JH, Park HE, Lee JC, Youn JI, et al. Seborrheic keratosis in the Korean males: causative role of sunlight. Photodermatol Photoimmunol Photomed. 2003;19(2):73-80.

29. Neel VA, Todorova K, Wang J, Kwon E, Kang M, Liu QS, et al. Sustained Akt Activity Is Required to Maintain Cell Viability in Seborrheic Keratosis, a Benign Epithelial Tumor. Journal of Investigative Dermatology. 2016;136(3):696-705.

30. Logie A, Dunois-Larde C, Rosty C, Levrel O, Blanche M, Ribeiro A, et al. Activating mutations of the tyrosine kinase receptor *FGFR3* are

associated with benign skin tumors in mice and humans. Hum Mol Genet. 2005;14(9):1153-60.

31. Hafner C, Hartmann A, Real FX, Hofstaedter F, Landthaler M, Vogt T. Spectrum of FGFR3 mutations in multiple intraindividual seborrheic keratoses. J Invest Dermatol. 2007;127(8):1883-5.

32. Hafner C, Lopez-Knowles E, Luis NM, Toll A, Baselga E, Fernandez-Casado A, et al. Oncogenic *PIK3CA* mutations occur in epidermal nevi and seborrheic keratoses with a characteristic mutation pattern. Proc Natl Acad Sci U S A. 2007;104(33):13450-4.

33. Webster MK, D'Avis PY, Robertson SC, Donoghue DJ. Profound ligand-independent kinase activation of fibroblast growth factor receptor 3 by the activation loop mutation responsible for a lethal skeletal dysplasia, thanatophoric dysplasia type II. Mol Cell Biol. 1996;16(8):4081-7.

34. Duperret EK, Oh SJ, McNeal A, Prouty SM, Ridky TW. Activating FGFR3 mutations cause mild hyperplasia in human skin, but are insufficient to drive benign or malignant skin tumors. Cell cycle. 2014;13(10):1551-9.

35. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204-13.

36. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. High burden and pervasive positive selection of somatic mutations in normal human skin. Science. 2015;348(6237):880-6.

37. Whitman M, Downes CP, Keeler M, Keller T, Cantley L. Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate. Nature. 1988;332(6165):644-6.

38. Madsen RR, Vanhaesebroeck B, Semple RK. Cancer-Associated *PIK3CA* Mutations in Overgrowth Disorders. Trends Mol Med. 2018;24(10):856-70.

39. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, et al. High frequency of mutations of the *PIK3CA* gene in human cancers. Science. 2004;304(5670):554.

40. Hart JR, Zhang YY, Liao LJ, Ueno L, Du LS, Jonkers M, et al. The butterfly effect in cancer: A single base mutation can remodel the cell. P Natl Acad Sci USA. 2015;112(4):1131-6.

41. Lim YH, Fisher JM, Bosenberg MW, Choate KA, Ko CJ. Keratoacanthoma Shares Driver Mutations with Cutaneous Squamous Cell Carcinoma. J Invest Dermatol. 2016;136(8):1737-41.

42. Mandinova A, Kolev V, Neel V, Hu B, Stonely W, Lieb J, et al. A positive *FGFR3/FOXN1* feedback loop underlies benign skin keratosis versus squamous cell carcinoma formation in humans. J Clin Invest. 2009;119(10):3127-37.

43. Seo EY, Lee DH, Lee Y, Cho KH, Eun HC, Chung JH. Microarray analysis reveals increased expression of DeltaNp63alpha in seborrhoeic keratosis. Br J Dermatol. 2012;166(2):337-42.

44. Cancer Genome Atlas Research N. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature. 2014;507(7492):315-22.

45. Seo J, Choi IH, Lee JS, Yoo Y, Kim NK, Choi M, et al. Rare cases of congenital arthrogryposis multiplex caused by novel recurrent *CHRNG* mutations. J Hum Genet. 2015;60(4):213-5.

46. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 2007;17(11):1665-74.

47. Diskin SJ, Li M, Hou C, Yang S, Glessner J, Hakonarson H, et al. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Res. 2008;36(19):e126.

48. Wang K, Chen Z, Tadesse MG, Glessner J, Grant SF, Hakonarson H, et al. Modeling genetic inheritance of copy number variations. Nucleic Acids Res. 2008;36(21):e138.

49. Sun BK, Boxer LD, Ransohoff JD, Siprashvili Z, Qu K, Lopez-Pajares V, et al. *CALML5* is a *ZNF750*-and *TINCR*-induced protein that binds

stratifin to regulate epidermal differentiation. Genes & development. 2015;29(21):2225-30.

50. Boxer LD, Barajas B, Tao S, Zhang J, Khavari PA. *ZNF750* interacts with *KLF4* and *RCOR1*, *KDM1A*, and *CTBP1/2* chromatin regulators to repress epidermal progenitor genes and induce differentiation genes. Genes & development. 2014;28(18):2013-26.

51. Sen GL, Boxer LD, Webster DE, Bussat RT, Qu K, Zarnegar BJ, et al. *ZNF750* Is a p63 Target Gene that Induces *KLF4* to Drive Terminal Epidermal Differentiation. Dev Cell. 2012;22(3):669-77.

52. Cohen I, Birnbaum RY, Leibson K, Taube R, Sivan S, Birk OS. *ZNF750* Is Expressed in Differentiated Keratinocytes and Regulates Epidermal Late Differentiation Genes. Plos One. 2012;7(8).

53. Birnbaum RY, Zvulunov A, Hallel-Halevy D, Cagnano E, Finer G, Ofir R, et al. Seborrhea-like dermatitis with psoriasiform elements caused by a mutation in *ZNF750*, encoding a putative C2H2 zinc finger protein. Nat Genet. 2006;38(7):749-51.

54. Sawada G, Niida A, Uchi R, Hirata H, Shimamura T, Suzuki Y, et al. Genomic Landscape of Esophageal Squamous Cell Carcinoma in a Japanese Population. Gastroenterology. 2016;150(5):1171-82.

55. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 2014;42(Database issue):D986-92.

56. Willatt L, Cox J, Barber J, Cabanas ED, Collins A, Donnai D, et al. 3q29 microdeletion syndrome: clinical and molecular characterization of a new syndrome. Am J Hum Genet. 2005;77(1):154-60.

57. Cox DM, Butler MG. A clinical case report and literature review of the 3q29 microdeletion syndrome. Clin Dysmorphol. 2015;24(3):89-94.

58. Martins FC, De S, Almendro V, Gonen M, Park SY, Blum JL, et al. Evolutionary pathways in *BRCA1*-associated breast tumors. Cancer Discov. 2012;2(6):503-11.

59. Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, Grinfeld J, et al. Effect of mutation order on myeloproliferative neoplasms. N Engl J Med. 2015;372(7):601-12.

60. Alves MM, Fuhler GM, Queiroz KC, Scholma J, Goorden S, Anink J, et al. *PAK2* is an effector of *TSC1/2* signaling independent of mTOR and a potential therapeutic target for Tuberous Sclerosis Complex. Sci Rep. 2015;5:14534.

61. Renkema GH, Pulkkinen K, Saksela K. Cdc42/Rac1-mediated activation primes *PAK2* for superactivation by tyrosine phosphorylation. Mol Cell Biol. 2002;22(19):6719-25.

62. Xuan Y, Chi L, Tian H, Cai W, Sun C, Wang T, et al. The activation of the NF-kappaB-JNK pathway is independent of the PI3K-Rac1-JNK pathway involved in the bFGF-regulated human fibroblast cell migration. J Dermatol Sci. 2016;82(1):28-37.

63. Appledorn DM, Dao KH, O'Reilly S, Maher VM, McCormick JJ. Rac1 and Cdc42 are regulators of HRasV12-transformation and angiogenic factors in human fibroblasts. BMC Cancer. 2010;10:13.

64. Heidenreich B, Denisova E, Rachakonda S, Sanmartin O, Dereani T, Hosen I, et al. Genetic alterations in seborrheic keratoses. Oncotarget. 2017;8(22):36639-49.

65. Siu MK, Kong DS, Ngai SY, Chan HY, Jiang L, Wong ES, et al. p21-Activated Kinases 1, 2 and 4 in Endometrial Cancers: Effects on Clinical Outcomes and Cell Proliferation. Plos One. 2015;10(7):e0133467.

66. Hao S, Luo C, Abukiwan A, Wang G, He J, Huang L, et al. miR-137 inhibits proliferation of melanoma cells by targeting *PAK2*. Exp Dermatol. 2015;24(12):947-52.

67. Gao C, Ma T, Pang L, Xie R. Activation of P21-activated protein kinase 2 is an independent prognostic predictor for patients with gastric cancer. Diagn Pathol. 2014;9:55.

68. South AP, Purdie KJ, Watt SA, Haldenby S, den Breems NY, Dimon M, et al. *NOTCH1* mutations occur early during cutaneous squamous cell carcinogenesis. J Invest Dermatol. 2014;134(10):2630-8.

69. Katoh M, Katoh M. FGF signaling network in the gastrointestinal tract (review). Int J Oncol. 2006;29(1):163-8.

70. Nowell CS, Radtke F. Notch as a tumour suppressor. Nat Rev Cancer. 2017;17(3):145-59.

71. Nicolas M, Wolfer A, Raj K, Kummer JA, Mill P, van Noort M, et al. Notch1 functions as a tumor suppressor in mouse skin. Nat Genet. 2003;33(3):416-21.

72. Maxwell KN, Wubbenhorst B, Wenz BM, De Sloover D, Pluta J, Emery L, et al. BRCA locus-specific loss of heterozygosity in germline *BRCA1* and *BRCA2* carriers. Nat Commun. 2017;8(1):319.

국 문 초 록

어떤 종류의 DNA 변화는 양성종양과 악성종양에서 공유되며, 같은 유전자 변화에도 불구하고 세포, 조직, 임상적인 경과가 다르다. 그렇기에 양성종양의 유전체학을 연구하는 방법으로 악성종양 치료의 새로운 접근법을 제시할 수 있을 것이다. 그러나 현재 양성 종양의 유전체는 많이 연구되지 않았다. 지루각화증은 피부의 양성종양이다. *FGFR3* 및 *PIK3CA* 변이가 지루각화증을 일으키는 것이 발견되었으나, 전장 유전체 분석은 아직 이루어진 바 없다. 본 연구는 49 명의 환자에게서 얻은 51 개의 지루각화증과 정상 침 샘플들을 분석하였다. 8 개의 조직에서 전장 엑솜 시퀀싱을 통해 Somatic mutation, copy number variation, loss of heterozygous 분석을 진행하였고, 나머지 43 개 샘플에서는 생어 시퀀싱과 SNP array 를 이용하여 전장 엑솜 시퀀싱에서 발견된 변이들의 분포를 연구하였다. 그 결과 *FGFR3*, *PIK3CA*, *ZNF750* 유전자들의 변이가 지루각화증의 major driver 임을 발견하였으며, 그 외에도 *NOTCH1*, chr3q29 microdeletion, *MSH2–MSH6*, *TERT1* promoter, *RB1* 등도 driver 로 작동할 수 있음을 발견했다. *FGFR3* 과 *ZNF750* 의 loss of function 조합은 지루각화증의 acanthotic subtype 과 연관되었고, 단 두 유전자의 변이만으로 지루각화증이 발생하는 것으로 보인다. *FGFR3* 과

PIK3CA activation 변이는 hyperkeratotic subtype 과 연관되어 있었다. Minor allele frequency 분석과 Laser microdissection 실험을 통해 *FGFR3*, *ZNF750*, *PIK3CA* 변이가 특정 순서를 따라 발생함을 보였다. Base change 분율 분석을 통해 자외선 노출 정도와 나이가 각각 지루각화증의 유전체학적인 risk factor 임을 보였다. 본 연구를 통해 지루각화증의 유전체학적 배경과 발생과정, 조직학적 분화의 유전적 원인을 제시할 수 있었다.

주요어: 전장 엑솜 분석, 지루각화증, *FGFR3*, *ZNF750*, *PIK3CA*, 암 진화

학 번: 2012-22262