

2015

Improving DNA Quality using FFPE tissues for Array Comparative Genomic Hybridization to find Single Nucleotide Polymorphisms (SNPs) in Melanoma

Keerti Potluri
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all



Part of the [Molecular Biology Commons](#)

Repository Citation

Potluri, Keerti, "Improving DNA Quality using FFPE tissues for Array Comparative Genomic Hybridization to find Single Nucleotide Polymorphisms (SNPs) in Melanoma" (2015). *Browse all Theses and Dissertations*. 1314.

https://corescholar.libraries.wright.edu/etd_all/1314

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.

Improving DNA quality using FFPE tissues for Array Comparative Genomic
Hybridization to find Single Nucleotide Polymorphisms (SNPs) in Melanoma

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science

By

KEERTI POTLURI

B.Tech., JNTU University, 2010

2015

Wright State University

WRIGHT STATE UNIVERSITY
GRADUATE SCHOOL

July 22, 2015

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Keerti Potluri ENTITLED Improving DNA quality using FFPE tissues for Array Comparative Genomic Hybridization to find Single Nucleotide Polymorphisms (SNPs) in Melanoma. BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Committee on Final
Examination

Michael Markey, Ph.D.
Associate Professor of
Biochemistry & Molecular
Biology

Michael Kent, Ph.D.
Translational Scientist
Dermopath Lab of Central
States (DLCS)

John Paietta, Ph.D.
Associate Professor of
Biochemistry & Molecular
Biology

Robert E.W. Fyffe, Ph.D.
Vice President for Research and
Dean of the Graduate School

Michael Markey, Ph.D.
Thesis Director

Madhavi P. Kadakia, PhD.
Chair, Department of Biochemistry and
Molecular Biology

ABSTRACT

Potluri, Keerti. M.S. Department of Biochemistry and Molecular Biology, Wright State University, 2015. ENTITLED Improving DNA quality using FFPE tissues for Array Comparative Genomic Hybridization to find Single Nucleotide Polymorphisms (SNPs) in Melanoma.

Important classes of benign and malignant neoplasms are composed of melanocytic cells which produce a pigment called melanin. Benign nevi, which are non-malignant melanocytic lesions, can sometimes give rise to malignant melanoma. Melanoma can be a lethal melanocytic neoplasm, a deadly and aggressive form of skin cancer. Finding prognostic or diagnostic markers can be very useful to reduce the deaths caused by melanoma. Array Comparative Genomic Hybridization (aCGH), a cytogenetic technique, analyzes the whole genome or chromosome for detecting genetic aberrations/variations such as single nucleotide polymorphisms (SNPs) in a cancer. formalin fixed paraffin embedded FFPE tissues are usually taken from suspected tumor tissues, fixed with formalin to preserve the protein and cytoskeletal structure and embedded in paraffin wax so that they can be cut and used for pathological diagnosis. These specimens are the starting material for extracting the DNA, but it can be quite challenging for getting DNA of good quality. Here we compared 27 samples extracted from FFPE tissues with three different extraction methods: phenol-chloroform isoamyl alcohol, Qiagen QIAamp FFPE

kit, and adaptive focused acoustics (AFA) to see whether the method will have any effect on the DNA quality in the samples. We found that the AFA method showed better quality control (QC) results than other methods, where AFA showed increased amplicon length and decreased RAPD PCR failure rate. These were successfully hybridized to the microarrays and the data compared between methods. A total of 42 melanocytic nevi and 21 benign nevi were analyzed by aCGH. We found some novel SNPs and the genes associated with them, these genes are already shown to be involved in melanoma as well as in benign nevi. So, these findings can help to see whether the SNPs from benign nevi are predisposing to melanoma or if the SNPs themselves are causing the melanoma and help identify potential therapeutic targets.

HYPOTHESIS

The hypothesis of the project is that the quality of the DNA does have an effect on how successfully the samples can hybridize onto array CGH microarrays. Therefore, improving the quality of the DNA will improve the aCGH data to efficiently find SNPs that cause/predispose to malignant melanoma, which can be useful for finding prognostic/diagnostic markers.

PROJECT AIMS

1. Optimize the quality of extracted genomic DNA from formalin-fixed paraffin embedded tissues for DNA fragment length, yield, and PCR performance.
2. Improve the performance of SNP 6.0 microarrays for FFPE DNA and generate raw array data.
3. Detect the novel SNPs that predispose to malignant melanoma.

TABLE OF CONTENTS

I.INTRODUCTION AND BACKGROUND.....	1
Melanoma.....	1
Survival rates.....	1
Current Diagnostic Techniques.....	3
Array Comparative Genomic Hybridization.....	5
Single Nucleotide Polymorphisms (SNPs) in Melanoma.....	6
Advantages of formalin fixed paraffin embedded (FFPE) tissues.....	7
Difficulties of extracting DNA from FFPE tissues.....	7
Genotyping analysis using Genotyping Console (GTC).....	9
Partek Genomics Suite (PGS) for filtering and analyzing SNPs.....	9
Summary of findings using Ingenuity Pathway Analysis (IPA).....	9

II. MATERIALS AND METHODS.....	10
Specimen collection and preparation.....	10
DNA isolation from FFPE tissues.....	11
Qiagen Extraction Method.....	11
Phenol chloroform Isoamyl Extraction Method.....	18
Adaptive Focused Acoustics (AFA) Extraction Method.....	21
Quality Control.....	25
Qubit.....	25
Electrophoresis.....	28
RAPD PCR.....	29
Microarrays.....	30
Data Analysis.....	32
Genotyping Console.....	33
gPLINK.....	34
Partek Genomics Suite.....	36

Ingenuity Pathway Analysis.....	40
III. RESULTS AND DISCUSSION.....	41
Requirements for collecting FFPE tissue specimens.....	41
H&E Staining of FFPE tissues.....	42
Comparing three extraction methods by looking at Quality Control measures...	43
Checking the Quality of the DNA in the SNP 6.0 Protocol	51
Microarray Analysis.....	55
Genotyping Analysis in the GTC.....	59
Analysis of SNPs in the PGS.....	65
Confirming and finding SNPs in IPA.....	69
III. CONCLUSION.....	78
IIII. REFERENCES	82

LIST OF FIGURES

1. H&E staining for the first and last slides of melanoma tissues.....	42
2. RAPD PCR comparison for the three extraction methods.....	46
3. DNA yield per section of three methods.....	48
4. A260/280 ratios comparisons of three extraction methods.....	49
5. A260/230 ratios comparisons of three extraction methods.....	49
6. Genomic DNA comparison of three extraction methods.....	50
7. Maximum Amplicon Length in RAPD PCR of three methods.....	51
8. Post PCR gel image in SNP 6.0 Protocol.....	52
9. Fragmentation of PCR products.....	54
10. Array CGH APT Fragment Length Analysis contrast QC comparison.....	58
11. Distribution of Contrast Values in CQC.....	60
12. Flow chart showing the QC steps for both Samples and SNPs in GTC.....	62
13. Flowchart showing the steps in getting the filtered SNPs using PGS and IPA to confirm the data.....	68
14. Flowchart showing sequential steps in analysis.....	77

LIST OF TABLES

1. Details of the patients used in the aim 1	42
2. Summary of the QC measures of three Methods Comparison.....	45
3. Comparison of factors for the three extraction methods.....	47
4. QC measures of PCR products of AFA samples in SNP 6.0 protocol...53	
5. Details of the 63 specimens.....	55
6. Summary of filtering out number of SNPs and Samples in each statistical tool.....	64
7. Statistical Tools with their parameters and thresholds.....	67
8. Genes and associated SNPs from IPA shown to be involved the melanoma.....	69
9. Genes and new SNPs found in our samples.....	73

I.INTRODUCTION AND BACKGROUND

Melanoma is one type of skin cancer that is serious. This cancer occurs in melanocyte cells which are present in the skin like uvea of the eye and also it can occur in mouth, anal or genital areas and other organs inside the body like the intestines. The melanocytes are present in one of the layers of skin called epidermis where they produce a pigment called “melanin”. This melanin gives color to the skin based on how much is present in the skin, the more the melanin the darker the skin will be. Of all the skin cancers melanoma is least common (2%), yet it is the cause of most skin cancer-related deaths [2]. According to the American Cancer Society, in 2015 approximately 74,000 cases will be diagnosed in the US and approximately 10,000 people will die from this cancer. Benign nevi which are noncancerous (moles) and non-metastatic when exposed to risk factors like UV light sometimes develop into cancerous tumors and spreads to the lymph nodes, blood vessels and finally to other organs like the lungs, brain etc. and become metastasized. It is a rare cancer and curable if detected in the early stages [3] but is deadly once metastasized everywhere. The 5 year survival rate shows that depending on the stage of the cancer when diagnosed, it can be as high as 98% when it does not spread, it decreases to 67% if it spreads to other lymph nodes and to less than 15% if it spreads to other organs [2]. People with risk factors like red hair, fair skin and blue eyes are at increased risk of melanoma [4] but these factors are only some of the genetic risk factors

associated with melanoma. Melanocytes can give rise to both benign melanocytic nevi and malignant melanoma. A high risk of melanoma is found in if the people having a family history of the disease [5]even though it is otherwise rare. Some studies have shown that those who had familial melanoma, and so many benign melanocytic nevi there is strong chance that the nevi can give rise to the melanoma[6-9].

Early diagnosis of melanoma is very important for treatment. Histopathologic study of skin biopsies by dermatopathologists is considered as the standard approach to diagnose skin lesions with the help of microscope [10]. Cellular anatomy of a tumor revealed by Hematoxylin and Eosin (H&E) staining which is one of the most widely used standard method to diagnose the suspected tumors. This staining method helps in showing staging and grading of the tumor taken from the skin biopsies. These are developed to determine the treatment and prognosis of malignant neoplasms. Staging tells the condition of cancer depending on whether the size or extension of invasion is large or not, if the invasion is spread to lymph nodes or not and tells the presence/absence of distant metastases. Stage 0 shows that the melanoma is in situ, stage I, II and III show that the tumor size is large and spread to other organs like lymph nodes and/or nearby organs to primary tumor site, stage III shows that the cancer has already spread to distant organs [11] On the other hand, grading is dependent upon the microscopic observation of the malignant neoplasms in H&E staining. It tells about the abnormality of tumor tissues and look under microscope. The higher the grading the less differentiation of cells and the worse the behavior of malignant neoplasms and vice versa, since benign neoplasms are well differentiated and differentiated cells have the resemblance of the original cells [12]. However, histopathology does not show the molecular changes. Since it is sometimes difficult to

differentiate benign nevi from malignant melanoma, because some melanocytic tumors share common histopathological features that are overlapping. Some benign melanocytic nevi, because of secondary changes, show abnormal characteristics that are mostly associated with the melanoma diagnosis. Therefore the histopathologic diagnosis of these tumors can be a challenge [13]. The histopathologic examination is sometimes not enough to make the diagnostic decisions. Several studies reported the uncertainty and discordance in diagnosing melanoma from benign nevi among the dermatopathologists who are experts in this area [13-19]. The ambiguity of melanoma can lead to misdiagnosis. Over diagnosis makes more stress, unnecessary surgeries and cost of medical insurance. On the other hand, under diagnosis can result in negligence of the disease[3] and can be deadly but curable if detected in the early stages. Because of this, molecular diagnostic techniques have become ancillary tools to aid diagnosis of melanocytic lesions. These techniques have shown to be useful in identifying the melanocytic lesions and the presence of melanoma.

BRAF is a serine/threonine protein kinase involved in the MEK/ERK pathway. Mutations in this protein; BRAF V600E (about 80%) and BRAF V600K (about 5-30%) are most prevalent in melanoma [20]. Vemurafenib is an inhibitor of mutant BRAF activity. Patients in late stage melanoma when treated with Vemurafenib have improved survival chances [21]. These mutations can occur later in life as oncogene, causing cancer and is seen in both melanoma and rarely in nevi. To know the existence of these mutations, histopathological observations are not enough. Molecular diagnostic methods like Immunohistochemistry (IHC) are available as adjunct techniques for diagnosing melanoma by looking into melanocytic lesions [22] [23, 24] by looking at the protein

expression level. Some of the markers used for IHC include: S-100, HMB-45, Mitf, Melan-A, and tyrosinase for diagnosing melanoma [25] and also Ki-67 which is used as mitotic index to determine the stage of the cell tumor [26]. S-100 protein is a calcium binding protein and it has association with malignant melanoma because the higher expression of it leads to cancer cell proliferation, differentiation, tumor metastases, cell invasion, cell growth and apoptosis [27]. S100 protein monoclonal antibody thus can help in diagnosing melanoma from tumors by reacting with the antigens in the melanocytic tumors with doubtful histological origin and tells whether the presence of tumors is in lymph nodes or other tissues [27] [28]. HMB-45 is a monoclonal antibody that reacts to antigens in the melanocytic tumors only. It has a good specificity for identifying malignant melanomas thus avoiding poor diagnosis [29, 30]. Mitf is a transcription factor involved in the regulation of pathways of cells like melanocytes, it regulates the production of melanin in melanocytes. Any mutation in Mitf leads to melanoma [31]. Recently Mitf immunohistochemical stain is used in identifying metastatic melanoma effectively [32]. Melan-A is a protein antigen found in the melanocytes. The melan-A marker is found to be useful in diagnosing melanoma [33] [34]. Ki-67 is a protein involved in cell proliferation and the marker can help in identifying the cell stage especially during mitosis to determine the growth of the tumor [26]. But all these markers still lack a certain group of lesions indeterminate even to expert pathologists. Additionally, cytogenetic techniques have emerged to diagnose melanocytic lesions. Fluorescence in situ hybridization (FISH) is one of these techniques that have gained use in melanoma diagnostics [10, 35-37]. FISH recently has a set of new four probes targeting 6p25 (RREB1), 6q23 (MYB), 11q13 (CCND1), and centromere 6 (CEP6) that

are being used to identify the malignant melanomas with increased specificity and sensitivity [13]. These 4 probe set were selected as best for distinguishing benign nevi from malignant melanoma [38, 39]. RREB1 also known as Ras-responsive element binding protein 1 (RRE) and Raf responsive zinc finger protein is a transcription factor that binds to RRE and leads to increased Ras/Raf activity. MYB is another transcription factor. CCND1 is cyclin D1 proto oncogene which plays an important role in G1-S phase. CEP6 is used as a control for the ploidy status for chromosome 6. When tested with these 4 genes in so many studies, they were able to distinguish benign from melanoma in most of the cases, thus this 4 probe panel became significant in diagnosing the melanoma. However FISH has its own limitations that it cannot detect smaller abnormalities in the chromosome, cannot be used for screening the entire chromosomes and is loci specific [40, 41]. Another cytogenetic technique, comparative genomic hybridization (CGH) has become a useful technique since it can look into the entire genome to look for abnormalities. CGH has been a very useful technique in cancer research [42] and in diagnosing melanoma [43]. Even though it has the ability to distinguish the tumors it has its own disadvantages of having very less resolution of 3-10 Mb [44] which means its inability to detect smaller regions in copy number changes and certain chromosomal regions involved in losses and gains [40] across the genome. So this lead to another technique called array CGH with higher resolution to look into the entire genome to analyze abnormalities on entire genome and/or chromosome of our interest. Array comparative genomic hybridization (aCGH) can detect simultaneously single nucleotide polymorphisms (SNP), is a DNA sequence variation occurring in which a single nucleotide (A, T, C or G) in the genome or other shared sequence differs between members of a species or

within the paired chromosomes. Copy number variations that have alterations of DNA across the genome with more than normal number of segments of DNA (gene) either deleted or amplified can also be detected in a single experiment. aCGH is a highly specific, fast, and high throughput technique. It can be used to distinguish between benign nevi and malignant melanoma by using FFPE tissues [45]. It can be used on a variety of sample types such as paraffin embedded [46], frozen [47], cell lines [48], blood samples [49] and has major applications in cancer [50]. It has shown promising results for diagnosing many cancers [50]. An important advantage is that it requires only a minimum number of cells from FFPE tissue sections [45].

Because of its high resolution, array CGH is being used to identify the most frequent genomic variants called SNP. SNPs can affect the function of genes and SNPs act as biological markers in finding the genes involved with diseases [51]. Studies have shown that the SNPs are useful in studying cancer [52, 53]. Benign nevi are associated with an increased risk of melanoma [54]. Therefore it is important to look into both benign and melanoma to find the actual SNPs associated with melanoma. Some of the genes that are associated with SNPs and are involved in melanoma are BRAF [55], GNAQ [56], GNA11 [57], PLA2G6, MTAP and IRF4[54]. We are interested in SNPs and genes that are associated with melanoma.

Array CGH has its own limitations. It requires good quality and a sufficient quantity of DNA to successfully hybridize to the microarrays. Getting good quality DNA from formalin fixed paraffin embedded tissues (FFPE) remains challenging. FFPE tissues are

used as starting material in our research. Because it has many advantages such as stability for decades [58], can be used for routine histopathologic diagnosis of diseases [58], easy handling and especially used for downstream analysis [59] At the same time DNA obtained from biopsied tissues are small and often irreplaceable. We have to extract DNA by an efficient extraction method. While FFPE tissues are stable for decades [60], this preservation presents significant challenges. The process of fixation with formalin, embedding and handling the tissues, should be done carefully in order to improve the quality of the nucleic acids for molecular testing [61]. Formalin treatment cross-links biological molecules such as DNA and proteins [61] so the nuclear material will get trapped and prevent from breakage which later has to be reverse cross linked to retrieve the nuclear material. Also, since longer nucleic acids like RNA and DNA, are fragmented in the preservation process, it leads to poor performance in downstream analyses [62, 63]. To optimize utility of nucleic acids from FFPE specimens for aCGH, we must reverse these cross links and avoid further degradation during the DNA extraction process. Inefficient tissue rehydration results in poor reverse cross linking and thereby less quality of DNA, so it is very important to rehydrate the tissue during paraffin removal. Several previous studies have compared protocols for DNA extraction from FFPE tissues, with varying results. Senguven et al. compared several variations on the CTAB method that used CTAB to disrupt the cellular membranes and a hot alkali method that lyse the genetic material with alkaline solution, and these both methods are compared with the Qiagen FFPE DNA kit [64], a solid-phase extraction that uses silica gel membrane in the spin column to trap the nucleic acids and eluted later with buffer. They found that the Qiagen columns performed best, and that DNA yield became especially low in samples

more than 50 years old. There is some controversy between laboratories, however, as other studies have compared the Qiagen kit with phenol-chloroform. This is a liquid-liquid extraction method separates the proteins in organic phase and DNA in the aqueous phase. They found either the columns [65] or the phenol method [66] to be superior in yield and amplifiability. Other methods less frequently used, not measured in our study include the Maxwell 16 method by Promega (Mannheim, Germany). In this method, the nucleic acids bound to magnetized silica particles and later eluted. This extraction method was shown to perform well against other automated Qiagen kits [67], and the Chelex bead method (Bio-Rad, Berkeley, California). This is an ion exchange resin that inactivates the nucleases by removing magnesium and thus protects DNA for the PCR. Chelex bead method has been used with some success in manual or automated extractions from FFPE tissues [68]. A small number of studies have shown that the DNA obtained by a recent sonication-based method called adaptive focused acoustics (AFA) is more amenable to next generation sequencing than DNA from other methods, resulting in greater coverage of the genome and facilitating easier assembly and alignment [69-71]. AFA is also a solid-phase extraction but it uses acoustic energy to de-paraffinize the tissues. To our knowledge, ours is the first study to compare AFA for use in aCGH. Here, we compared three different extraction methods: phenol-chloroform extraction, commercial (Qiagen) column-based purification, and adaptive focused acoustic (AFA) extraction, to see which works best to retrieve DNA for aCGH. Quantity and quality measures are considered for each extraction method. Nanodrop concentrations which are based on spectrophotometric absorbance at 260 nm are measured and also the 260/230 and 260/280 ratios were measured to check the purity of the samples. Since Nanodrop

concentrations are not accurate because the absorbance at 260 nm can measure not just dsDNA, so we measured qubit concentrations, since fluorescent dye in qubit only binds to the dsDNA and gives accurate concentrations of our samples. Next, we performed RAPD PCR that amplifies DNA template randomly to see whether the DNA in our samples are degraded or not by doing gel electrophoresis. Finally we measured RAPD PCR failure rates in each method, samples that have not shown enough amplicon length (< 300base pair length) are considered as failure and will be removed.

Once we got good quality of DNA then samples were hybridized them on to the microarrays successfully. These microarray data were analyzed using Genotyping Console (GTC) for genotyping analysis to get SNPs list. Then the SNPs were filtered and analyzed using Partek Genomics Suite (PGS). Once got the SNPs and associated genes list, used Ingenuity Pathway Analysis (IPA) to confirm our findings with the curated literature in IPA about the genes and SNPs known to be involved with melanoma. We found total of 8 genes along with the associated SNPs. Along with 8 genes known to be involved in melanoma, their associated novel SNPs were identified in genes that are known to be involved in melanoma.

II. METHODS AND MATERIALS

Specimen Collection and preparation:

After IRB review, the FFPE tissue blocks were collected from a national dermatopathology laboratory (Dermatopathology Laboratory of Central States, DLCS, Dayton, OH). They were stored in a temperature-controlled room. First we retrieved the clinical data and health records of the patients using laboratory information management software (Intellipath) at DLCS. The biopsy specimens collected are from 2007 and 2004, making them 8 or 11 years old. Patients ranged in age from 16 to 75 years. Based on the clinical data, we selected specimens that are unambiguously benign or melanoma and of sufficient size to not exhaust the tissue block during analysis.

10 μ m thickness sections were taken from FFPE tissue blocks using a microtome with disposable blades at DLCS. To avoid contamination between blocks, care was taken by wearing gloves and using new blades for every specimen. Sections were slightly warmed in the water bath to facilitate mounting onto slides. The slides were then air dried overnight. The first and last slides were kept for H&E staining and examined under microscope to confirm that the sections on the slides have the region of interest (benign melanocytic nevi or melanoma cellular material). The tissues in the sections were scrapped off avoiding as much paraffin as possible using sterile scalpel blades and collected into 1.5 ml centrifuge tubes for extracting the DNA. The number of sections and the amount of tissue taken were dependent on the extraction method (see below).

DNA isolation from FFPE tissues:

For DNA isolation three extraction methods were used: Qiagen QIAamp DNA FFPE tissue kit, phenol-chloroform isoamyl alcohol extraction and Adaptive Focused Acoustics (AFA) - based extraction using the Covaris truXTRAC FFPE DNA kit.

Qiagen method:

The protocol suggested by Qiagen QIAamp DNA FFPE tissue kit (QIAGEN GmbH, D-40724 Hilden) is as follows:

1. Using a scalpel, trim excess paraffin off the sample block.
2. Cut up to 8 sections 5–10 μm thick. If the sample surface has been exposed to air, discard the first 2–3 sections.
3. Immediately place the sections in a 1.5 or 2 ml micro centrifuge tube and add 1 ml xylene to the sample. Close the lid and vortex vigorously for 10 sec.
4. Centrifuge at full speed for 2 min at room temperature (15–25°C).
5. Remove the supernatant by pipetting. Do not remove any of the pellet.
6. Add 1 ml ethanol (96–100%) to the pellet, and mix by vortexing. The ethanol extracts residual xylene from the sample.
7. Centrifuge at full speed for 2 min at room temperature.

8. Remove the supernatant by pipetting. Do not remove any of the pellet. Carefully remove any residual ethanol using a fine pipet tip.

9. Open the tube and incubate at room temperature or up to 37°C. Incubate for 10min or until all residual ethanol has evaporated.

10. Resuspend the pellet in 180 µl Buffer ATL. Add 20 µl proteinase K, and mix by vortexing.

11. Incubate at 56°C for 1 h (or until the sample has been completely lysed).

12. Incubate at 90°C for 1 h. The incubation at 90°C in Buffer ATL partially reverses formaldehyde modification of nucleic acids. Longer incubation times or higher incubation temperatures may result in more fragmented DNA. If using only one heating block, leave the sample at room temperature after the 56°C incubation until the heating block has reached 90°C.

13. Briefly centrifuge the 1.5 ml tube to remove drops from the inside of the lid. If RNA-free genomic DNA is required, add 2 µl RNase A (100 mg/ml) and incubate for 2 min at room temperature before continuing with step 14. Allow the sample to cool to room temperature before adding RNase A.

14. Add 200 µl Buffer AL to the sample, and mix thoroughly by vortexing. Then add 200 µl ethanol (96–100%), and mix again thoroughly

by vortexing. It is essential that the sample, Buffer AL, and ethanol are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. Buffer AL and ethanol can be premixed and added together in one step to save time when processing multiple samples. A white precipitate may form on addition of Buffer AL and ethanol. This precipitate does not interfere with the QIAamp procedure.

15. Briefly centrifuge the 1.5 ml tube to remove drops from the inside of the lid.

16. Carefully transfer the entire lysate to the QIAamp MinElute column (in a 2 ml collection tube) without wetting the rim, close the lid, and centrifuge at 6000 x g (8000rpm) for 1min. Place the QIAamp MinElute column in a clean 2ml collection tube, and discard the collection tube containing the flow-through. If the lysate has not completely passed through the membrane after centrifugation, centrifuge again at a higher speed until the QIAamp MinElute column is empty.

17. Carefully open the QIAamp MinElute column and add 500 µl Buffer AW1 without wetting the rim. Close the lid and centrifuge at 6000 x g (8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2ml collection tube, and discard the collection tube containing the flow-through.

18. Carefully open the QIAamp MinElute column and add 500 µl Buffer AW2 without wetting the rim. Close the lid and centrifuge at 6000 x g

(8000 rpm) for 1 min. Place the QIAamp MinElute column in a clean 2ml collection tube, and discard the collection tube containing the flow-through. Contact between the QIAamp MinElute column and the flow-through should be avoided. Some centrifuge rotors may vibrate upon deceleration, resulting in the flow-through, which contains ethanol, coming into contact with the QIAamp MinElute column. Take care when removing the QIAamp MinElute column and collection tube from the rotor, so that low-through does not come into contact with the QIAamp MinElute column.

19. Centrifuge at full speed (20,000 x g; 14,000 rpm) for 3 min to dry the membrane completely. This step is necessary, since ethanol carryover into the eluate may interfere with some downstream applications.

20. Place the QIAamp MinElute column in a clean 1.5 ml micro centrifuge tube (not provided), and discard the collection tube containing the flow-through. Carefully open the lid of the QIAamp MinElute column and apply 20–100 μ l Buffer ATE to the center of the membrane.

Important: Ensure that Buffer ATE is equilibrated to room temperature. If using small elution volumes (<50 μ l), dispense Buffer ATE onto the center of the membrane to ensure complete elution of bound DNA. QIAamp MinElute columns provide flexibility in the choice of elution volume. Choose a volume according to the requirements of the

downstream application. The volume of elute will be up to 5 μ l less than the volume of elution solution applied to the column.

21. Close the lid and incubate at room temperature for 1 min. Centrifuge at full speed (20,000 x g; 14,000 rpm) for 1 min, incubating the QIAamp MinElute column loaded with Buffer ATE for 5min at room temperature before centrifugation generally increases DNA yield.

The original protocol for the Qiagen FFPE DNA Tissue kit, modified by previous student in the lab (Sameep Naik), is as follows:

1. Took 10 μ m thickness sections and Incubate tissue sample at 60°C in water bath for 30 minutes.
2. Wash the tissue in micro centrifuge tube with 1 ml of 100% xylene and vortex, centrifuge at 20,000xg for 5 minutes, and then remove supernatant.
3. Wash tissue in 500 μ l of 100% ethanol, vortex and centrifuge at 20,000xg for 3 minutes then remove supernatant.
4. Wash tissue in 500 μ l of 75% ethanol, vortex and centrifuge at 20,000xg for 3 minutes then remove supernatant.
5. Wash tissue in 500 μ l of 50% ethanol, vortex and centrifuge at 20,000xg for 3 minutes and then remove supernatant.
6. Allow tissue to dry before processing.
7. Add 300 μ l buffer ATL to dried tissue.

8. Add 100µl proteinase K, mix by vortexing.
9. Incubate overnight (8-12 hours, never longer than 16 hours) at 56°C.
10. Add 400µl buffer AL to the sample and mix by inversion.
11. Incubate at 70°C for 10 minutes.
12. Add 400µl ethanol to the sample and mix by inversion.
13. Pipet mixture into DNeasy Mini Spin Column then centrifuge at 6000xg for 1 minute.
Discard the flow through.
14. Add 500µl wash buffer AW1, then centrifuge at 6000xg for 1 minute. Discard the flow through.
15. Add 500µl wash buffer AW2 then centrifuge at 6000xg for 1 minute. Discard the flow through.
16. Add 500µl wash buffer AW2 again, then centrifuge at 6000xg for 1 minute. Discard the flow through.
17. Replace flow through tube with clean tube, then centrifuge at 20,000xg for 3 minutes to remove any residual ethanol from the spin column.
18. Replace flow through tube with the clean micro centrifuge tube.
19. Add 100µl buffer ATE, incubate at room temperature for 5 minutes.
20. Centrifuge at 20,000xg for 1 minute.

We used the protocol of previous student but with some modifications as follows:

1. Cut 24 FFPE sections of 10 μm thickness for each DNA extraction.
2. Incubate tissue samples in water bath at 60°C for 30 minutes.
3. Wash tissues in micro centrifuge tube twice in 1 mL xylene, vortex and centrifuge at 20,000xg for 3 minutes each time.
4. Wash tissues in a descending concentration of ethanol (100, 75%, and then 50%), each time vortex and centrifuge at 20,000xg for 3 minutes and then remove supernatant.
5. Allow tissues to dry in the micro centrifuge tubes completely before proceeding to next step.
6. Add 300 μL Qiagen buffer ATL plus 40 μL proteinase K (20mg/mL, 5 PRIME Inc., Gaithersburg, MD, USA) to dried tissue and mix by vortexing, incubate overnight for first 24 hours.
7. Then add an additional 30 μL proteinase K at 24 hours and another 30 μL at 48 hours.
8. After 72 hours digestion, add 400 μl buffer AL to the sample, mix by inversion (never vortex after overnight incubation).
9. Incubate at 70°C for 10 minutes.
10. Add 400 μl ethanol to the sample and mix by inversion.
11. Pipet sample mixture into were Qiagen DNeasy Mini Spin Columns, then centrifuge at 6000xg for 1 minute. Discard the flow through.
12. Add 500 μl wash buffers AW1 and AW2 one by one, each time centrifuge at 6000g for 1 minute and discard the flow through.
13. Add 500 μl wash buffer AW2 again, then centrifuge at 6000xg for 1 minute. Discard the flow through.

14. Replace flow through tube with clean tube, then centrifuge at 20,000xg for 3 minutes to remove any residual ethanol from the spin column.
15. Replace flow through tube with the clean micro centrifuge tube.
16. Add 100µl buffer ATE, incubate at room temperature for 5 minutes.
17. Centrifuge at 20,000xg for 1 minute.

Phenol-chloroform-isoamyl alcohol extraction method:

Phenol-chloroform-isoamyl alcohol 25:24:1 (Fisher Scientific, Fair Lawn, New Jersey, USA) was used as described in Isola et al. The protocol is as follows:

1. Twenty to thirty 5µ- sections were deparaffinized in eppendorf tubes (2 x 1 ml xylene for 10 minutes each and 2 x 1 ml 100% ethanol for 10 minutes each).
2. After air drying at room temperature, samples were suspended in 1 ml DNA extraction buffer (0.3 mg/ml proteinase K (Sigma, St. Louis, MO), 100 mmol/L NaCl, 10 mmol/L Tris-HCl pH 8, 25 mmol/L EDTA pH 8, and 0.5% sodium dodecyl sulfate) and were incubated with shaking at 55°C overnight.
3. Additional proteinase-K (10µl from 20 mg/ml stock solution) was added 24 hours and 48 hours later for a total incubation time of 72 hours.
4. A 500µl sample mixed with 500 µl phenol chloroform isoamyl (Amresco, Solon, OH) was incubated at room temperature for 10 minutes and centrifuged.

5. DNA in the top layer was collected and precipitated with 250 μ l of 7.5 mol/L ammonium acetate and 1 ml of ice-cold 100% ethanol.
6. DNA was pelleted by centrifugation (14,000 rpm for 20 minutes).
7. Glycogen (0.1 mg/ml; Sigma) was added before centrifugation as a carrier to increase the volume of the pellet.
8. DNA was dissolved overnight in 20 to 40 μ l of TE buffer (10 mmol/L Tris, 1 mmol/L EDTA).

We used the same protocol but with some modifications:

1. 24 sections of 10 μ m thickness were used from each FFPE sample. FFPE sections were scraped off of the air-dried slides using sterile scalpel blades and collected into 1.5 mL micro centrifuge tubes.
2. Add 1ml of xylene to each tube, vortex and incubate for 10 min, centrifuge for 5 min at 20,000 x g. Discard the supernatant.
3. Repeat the above step again.
4. Add 1ml of 100% ethanol, vortex, incubate for 10 min, centrifuge for 3 min at 20,000 x g. Discard the supernatant.
5. Repeat the above step.
6. Add 1ml of 75% ethanol, vortex, incubate for 10 min, centrifuge for 3 min at 20,000 x g. Discard the supernatant.
7. Air dry the tubes at room temperature by inverting them on a tissue.
8. Add 985 μ l of DNA extraction buffer and 15 μ l of proteinase k.

9. Keep tubes now on water bath shaking incubator at 57°C (control temperature) or at 70°C (Stat temperature). Adjust the speed to 3-3½ rpm.
10. Keep checking the water level in the shaker regularly.
11. Add 10µl proteinase k to each tube after 24 hours and vortex very lightly.
12. Add 10µl proteinase k to each tube after 48 hours and vortex very lightly.
13. After 72 hours take out the tubes and keep them at 95°C for 40 min and bring them to room temperature.
14. Add 500µl of each tube and 500µl of phenol chloroform to one tube (shake the chloroform) and vortex.
15. Centrifuge at 20,000g for 1 min.
16. 2 layers will be formed in each tube, flip each tube for 5 sec and let them sit for 5 min at room temperature.
17. Centrifuge at 20,000rcf for 10 min at 4°C.
18. Transfer the top aqueous layer that has DNA to new tube.
19. Add chloroform approximately the amount of DNA the new tube has (for 100ul of DNA amount add 100ul of chloroform), flip the tube, incubate for 5 min at room temperature and centrifuge at 20,000rcf for 10 min.
20. Using pipette take put the top layer, put them in new tube. Discard the organic solvent by dumping them in the container in SASH.
21. Add 1ml of ice cold ethanol to each tube and then add 250µl of ammonium acetate
22. Keep the tubes in -80°C for about an hour and centrifuge at 14,000 rpm

DNA Purification:

1. Wash the pellet by adding 500µl 70% ethanol, invert several times, and spin at maximum speed at 4°C for 5 min.
2. Aspirate carefully. Dry the pellet completely(~5-10min)
3. Pre-warm the TE buffer (Tris 10Mm, EDTA 1Mm pH8.0) at 55°C and add 40µl buffer to each tube.
4. Incubate at room temperature for overnight.

Adaptive focused acoustics:

The third extraction method is adaptive focused acoustics (AFA) technology. This was performed using FFPE tissues with 10 µm thickness for 8-10 sections. The extraction was performed according to the protocol suggested by Covaris (Woburn, MA, USA) in the truXTRAC FFPE DNA kit. Slides were warmed on a heat block to 37 °C for 30 seconds. FFPE tissue was then scraped from the slides, avoiding paraffin, using Covaris SectionPicks. Sections were collected into Screw-Cap microTUBES by using FFPE SectionPicks provided by Covaris. AFA was performed per manufacturer's instructions ("protocol C") on a Covaris M220 Focused-Ultrasonicator. The protocol is as follows:

1. Open microTUBE Screw-Cap, add 100 µl Tissue SDS Buffer into microTUBE and load FFPE tissue (section or core). Affix Screw-Cap back in place.
2. Process the sample using the settings provided in the protocol to dissociate the paraffin while simultaneously rehydrating the tissue. During the AFA

process it is normal for the solution to turn milky white as the paraffin is emulsified.

System	Duty Factor	Peak Incident Power	Cycles per burst	Treatment Time	Temperature (Instrument)
S220 or E220	10%	175 Watts	200	300 sec	20 °C
S2 or E210	10%	5 (Intensity)	200	300 sec	20 °C
M220	20%	75 Watts	200	300 sec	20 °C
LE220	30%	450 Watts⁽¹⁾	200	300 sec	20 °C

Paraffin removal and tissue rehydration settings

System	Duty Factor	Peak Incident Power	Cycles per burst	Time	Temperature (Instrument)
S220 or E220	10%	175 Watts	200	10 sec	20 °C
S2 or E210	10%	5 (Intensity)	200	10 sec	20 °C
M220	20%	75 Watts	200	10 sec	20 °C
LE220	30%	450 Watts⁽¹⁾	200	10 Sec	20 °C

Proteinase K mixing settings

- Open Screw-Cap microTUBE, add 20 µl of Proteinase K solution to the sample and affix Screw-Cap back in place.
- Process the sample using the settings provided to properly mix Proteinase K with the sample.
- Protein digestion at 56°C. Insert the required number of Heat Block microTUBE Adapters into a Heat Block and set the temperature to 56°C.
- Load the microTUBE into the adapter once the heat block has reached its set point.
- An incubation time of 1 hour at 56°C is sufficient for sections 10 µm or less in thickness; 12-hour (i.e. overnight) incubation should be used for larger samples, such as 25 µm sections and cores. If the digestion is incomplete after 12 hours, add 20 µl of Proteinase K solution, mix, and

incubate for 1 more hour. Here homogenized tissue was digested for 2 hours instead of 1 hour in proteinase K provided by Covaris kit.

8. Incubate the samples at 80°C for 1 hour to reverse formaldehyde crosslinks.
9. Insert the required number of Heat Block microTUBE Adapters into a Heat Block and set the temperature to 80°C. Load the microTUBE into the adapter once the heat block has reached its set point.
10. If using the same heat block for both the 56°C & 80°C incubations, the microTUBE should be stored at room temperature until the heat block reaches 80°C.
11. Transfer the sample to a clean 1.5 ml micro centrifuge tube.

Optional: The sample can be treated with RNase A to remove RNA before DNA purification. Add 5µl of RNase A solution and incubate for 5 minutes at room temperature.

Then DNA purification is as follows:

Set heat block to 70°C and preheat the required volume of Buffer BE in a 1.5mL microfuge tube: (number of samples x 100 µl x 1.1)

1. Add 140 µl Buffer B1 to your sample and vortex thoroughly.
2. Add 160 µl ethanol (>96%) to the sample and vortex thoroughly.
3. Centrifuge at 10,000 x g for 2 minutes at room temperature. After centrifugation much of the paraffin will have formed a white layer, floating on top of the liquid.
4. Place a Purification Column into a provided Collection Tube.

5. While holding the sample tube at about the same angle as in the rotor, use a pipette to slowly recover the liquid layer, and transfer to the column. Transfer of a small amount of paraffin particles to the column is acceptable and will not interfere with the DNA purification.
6. Spin the assembly at 11,000 x g for 1 minute.
7. Discard the flow-through and place the Column back in the Collection Tube.
8. 1 st wash: Add 500 μ l Buffer BW. Spin the assembly at 11,000 x g for 1 minute.
9. Discard the flow-through and place the Column back in the Collection Tube.
10. 2 nd wash: Add 600 μ l Buffer B5. Spin the assembly at 11,000 x g for 1 minute.
11. Discard the flow-through and place the column in a new Collection Tube
12. Dry column: Spin the assembly at 11,000 x g for 1 minute.
13. Elute DNA - 1 st step: Place the Purification Column into a new 1.5 ml microfuge tube and add 50 μ l pre-warmed Buffer BE (70 °C) to the center of the column. Incubate at room temperature for 3 minutes. Spin the assembly at 11,000 x g for 1 minute.
14. Elute DNA – 2 nd step: Add a second aliquot of 50 μ l pre-warmed Buffer BE. Incubate again at room temperature for 3 minutes. Spin the assembly at 11,000 x g for 1 minute.

15. DNA is eluted in 100 µl Buffer BE.

We followed the same protocol but we changed the paraffin removal treatment time from 300sec to 510sec .For this extraction method the DNA was concentrated by speedvac without heat in order to get enough concentration for the SNP 6.0 protocol.

Quality control:

A Nanodrop spectrophotometer was used to quantify DNA concentration as well as determine the A^{260}/A^{230} and A^{260}/A^{280} ratios. Clean the Nanodrop first by 1.5 µl water and then by 1.5 µl of buffer, the type of buffer depends on which extraction method the DNA sample was eluted. Now load 1.5 µl of DNA sample on to the Nanodrop to check the concentration and ratios.

Qubit:

For a more accurate quantitation, the Qubit® dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA) was used to check the concentrations of dsDNA for all samples.

We followed the manufacturer's protocol. The protocol is as follows:

1. Set up the required number of 0.5-mL tubes for standards and samples.

The Qubit® dsDNA BR Assay requires 2 standards.

Note: Use only thin-wall, clear, 0.5-mL PCR tubes. Acceptable tubes include Qubit® assay tubes.

2. Label the tube lids. Note: Do not label the side of the tube as this could interfere with the sample read. Label the lid of each standard tube correctly. Calibration of the Qubit® fluorometer requires the standards to be inserted into the instrument in the right order.

3. Prepare the Qubit® working solution by diluting the Qubit® dsDNA BR Reagent 1:200 in Qubit® dsDNA BR Buffer. Use a clean plastic tube each time you prepare Qubit® working solution. Do not mix the working solution in a glass container. Note: The final volume in each tube must be 200 μL . Each standard tube requires 190 μL of Qubit® working solution, and each sample tube requires anywhere from 180–199 μL . Prepare sufficient Qubit® working solution to accommodate all standards and samples. For example, for 8 samples, prepare enough working solution for the samples and 2 standards: ~200 μL per tube in 10 tubes yields 2 mL of working solution (10 μL of Qubit® reagent plus 1990 μL of Qubit® buffer).

4. Add 190 μL of Qubit® working solution to each of the tubes used for standards.

5. Add 10 μL of each Qubit® standard to the appropriate tube, then mix by vortexing 2–3 seconds. Be careful not to create bubbles. Note: Careful pipetting is critical to ensure that exactly 10 μL of each Qubit® standard is added to 190 μL of Qubit® working solution.

6. Add Qubit® working solution to individual assay tubes so that the final volume in each tube after adding sample is 200 µL. Note: Your sample can be anywhere from 1–20 µL. Add a corresponding volume of Qubit® working solution to each assay tube: anywhere from 180–199 µL.

7. Add each sample to the assay tubes containing the correct volume of Qubit® working solution, then mix by vortexing 2–3 seconds. The final volume in each tube should be 200 µL.

8. Allow all tubes to incubate at room temperature for 2 minutes.

Reading standards and samples procedure are as follows:

1. On the Home screen of the Qubit® 2.0 Fluorometer, press DNA, then select dsDNA Broad Range as the assay type. The “Standards” screen is displayed. Note: If you have already performed a calibration for the selected assay, the instrument prompts you to choose between reading new standards and running samples using the previous calibration. If you want to use the previous calibration, press No and skip to step 5. Otherwise, continue with step 2.

2. On the Standards screen, press Yes to read the standards.

3. Insert the tube containing Standard #1 into the sample chamber, close the lid, then press Read. When the reading is complete (~3 seconds), remove Standard #1.

4. Insert the tube containing Standard #2 into the sample chamber, close the lid, then press Read. When the reading is complete, remove Standard #2. When the calibration is complete, the instrument displays the Sample screen.

5. Insert a sample tube into the sample chamber, close the lid, then press Read. When the reading is complete (~3 seconds), remove the sample tube. The instrument displays the results on the Sample screen. The value displayed corresponds to the concentration after your sample was diluted into the assay tube. To find the concentration of your original sample, you can record this value and perform the calculation yourself or the instrument can perform this calculation for you. Repeat step 5 until all samples have been read.

Electrophoresis:

Genomic DNA fragment sizes were first estimated by agarose gel electrophoresis of 250 ng DNA using 1 % agarose gels (90 mM Tris-borate, 2 mM EDTA, 1% agarose). Samples were stained using SYBR safe (Life Technologies) and ran at ~120 volts for about an hour and visualized under GE ImageQuant LAS-3000 camera (GE Healthcare Life Sciences, Piscataway, NJ, USA). The Jurkat genomic DNA (Thermo Scientific) was used as a positive control for comparison and Lambda DNA-HINDIII Digest (New England Biolabs) ladder which is 23,130bp was used for looking and comparing the size of the genomic DNA.

RAPD PCR:

Samples with visible DNA fragments as large as 23,000 base pairs (bp) were processed further by randomly amplified polymorphic DNA PCR (RAPD-PCR) to determine which sample is ideal for SNP 6.0 by looking at the amplifiability of each sample that have high molecular weight amplicons. Non-specific primers and PCR conditions were used to produce multiple amplicons from each sample.

RAPD-PCR reactions were carried out in a 20 μ L volume containing 25 ng DNA and using 10 μ l of GoTaq 2X Green Master Mix (Promega, Madison, WI USA) that contain bacterially derived Taq DNA polymerase, reaction buffers at optimal concentrations for efficient amplification of DNA templates by PCR, dNTPs, MgCl₂ and Go Taq® Green Master Mix contains two dyes (blue and yellow) that allow monitoring of progress during electrophoresis. PCR was performed in 0.2mL tubes in a GeneAmp PCR System 9700 thermocycler (Life Technologies, Carlsbad, CA, USA).

Primers used for RAPD-PCR were generated by Eurofins MWG Operon Inc (Huntsville, AL, USA). Sequences for the primer pairs and cycling parameters were as follows: 5'-AATCGGGCTG-3' and 5'-GAAACGGGTG-3', denaturation for 94°C for 2.5 minutes, then 45 cycles of amplification (1 minute at 94°C, 1 minute at 55°C and 2 minutes at 72°C) then final extension for 7 minutes at 72°C and holding at 4°C; or 5'-TGTGCCCAAGTGAAGACTCAG-3' and 5'-GAGTGAGCGGAGAGGGAACT-3', 45 cycles of 94° C denaturation for 1 minute, 35° C amplification for 1 minute, and 72° C extension for 2 minute and holding at 4° C . PCR products were resolved on 3% TBE agarose plus SYBR Safe dye (Life Technologies). In order to see which DNA extraction

method will give enough amount of DNA which is around 500bp amplicon length for SNP 6.0 protocol, gels were imaged with a GE ImageQuant LAS-3000 camera (GE Healthcare Life Sciences, Piscataway, NJ, USA) and visualization was carried out as described under Electrophoresis above.

Microarrays:

A larger set of 63 skin biopsy specimens (21 benign and 42 melanoma) that includes a subset of nine extracted by AFA method and the remaining were extracted by Qiagen method were first quantified by Nanodrop, Qubit and RAPD PCR, then processed and hybridized to Affymetrix SNP6.0 microarrays. 0.5 µg of genomic DNA was processed using the SNP 6.0 protocol and microarrays suggested by Affymetrix (Affymetrix, Santa Clara, CA, USA) with some modifications to the standard protocol.

The original protocol is in Affymetrix Genome-Wide Human SNP Nsp/Sty 6.0 User Guide which is as follows;

Day 1

1. Sty1 Digestion
2. Sty 1 ligation
3. Sty 1 PCR

Day 2

1. Nsp 1 digestion
2. Nsp 1 ligation
3. Nsp 1 PCR

Day 3

1. PCR QC 2% Agarose Gel
2. PCR Purification by Iso propanol method

Day 4

1. Fragmentation
2. Fragmentation QC 4% Agarose gel
3. Labelling

Day 5

1. Hybridization of genomic DNA to microarray chips

Day 6

1. Fluidics wash of microarray chips
2. Scan the microarray chips

Some modifications done by previous student (Sameep Naik) are as follows:

The input amount of DNA was increased from 250ng per restriction enzyme (Nsp1 and Sty1) to 500ng each. The number of PCR reactions were doubled from the suggested 3 for Sty1 and 4 for Nsp1 to 6 for Sty1 and 8 for Nsp1. It is important to note that the number of reactions was increased; the number of cycles in each reaction remained the same. The additional PCR reactions were combined as in the standard protocol. PCR cleanup was performed using isopropanol extraction (refer to Affymetrix User Bulletin 2: Improvements to step 7 of the SNP Assay 6.0, PCR cleanup, using an isopropanol

precipitation method, P/N 702968 Rev. 1). Hybridization, washing, staining and lastly scanning of the arrays which are controlled by Affymetrix Genechip Command Console (AGCC) software, these steps were followed by the manufacturer's protocol. Microarray data (CEL, ARR and CHP files) are obtained for every sample when microarray chips were scanned using the GeneChip Scanner 3000 7G using AGCC software

Data Analysis:

27 benign melanocytic lesions were extracted with each extraction method and these three methods were compared and determined statistically by looking into Nanodrop concentration (ng/section), 260/280 and 260/230 ratios, Qubit concentration (ng/section), RAPD-PCR amplicon lengths, and PCR failure rates using a two-tailed paired student's t-test of the mean and standard error.

The CEL files obtained were processed in Affymetrix Power Tools (APT) provided by Affymetrix. Contrast Quality Control (CQC) which is the primary QC for Genome wide Human SNP Array 6.0 was calculated and compared to restriction fragment sizes (for Nsp1 and Sty1) to evaluate microarrays to see which PCR products were successfully hybridized to the array. It measures the SNP signals and estimate how well they resolve into three genotype clusters. A subset of probes were used to measure the differences in contrast distributions for both homozygote and heterozygote genotypes. CQC will assign a call rate to each fragment length depending on how successfully they hybridized to the array. Optimally a CQC value ≥ 0.4 [72] is considered a successful hybridization of genomic DNA fragment to the arrays.

Genotyping Console Software:

The CEL files of 63 specimens (21 benign and 42 melanoma samples) were added in the genotyping console and performed primary intensity quality control (QC) for each sample. Intensity QC contains QC metrics, In/Out bounds, #CEL/CHP files, ARR files of all samples. First, samples were filtered based on CQC that has a default threshold value of greater than or equal to 0.4 which is showed by bounds whether the samples met QC value or not. Only samples that passed this QC should be used for genotyping analysis. The genotyping console uses algorithm called BIRDSEED V2, it is a SNP genotyping algorithm that runs on the Affymetrix SNP 6.0 platform. Birdseed uses a customized Expectation-Maximization (EM) algorithm to fit two-dimensional Gaussians to SNP data, producing genotypes and confidence scores for every individual at every SNP. Since it is a clustering algorithm, it runs on so many samples at the same time (50 or more). The batch genotyping results that has sample QC metrics like CQC will be shown in CHP summary table. This table generates the batch CHP files. These output format files of genotyping console can be used as input files in downstream applications like Partek Genomics Suite (PGS) for further analysis. With the help of CHP summary table, SNPs list can be generated by filtering the QC call rate with a default threshold of $\geq 90\%$. The call rate is defined as the fraction of called SNPs per sample over the total number of SNPs in the dataset. Any samples with SNP call rates below this threshold value are omitted from the further analysis since these can cause false positive results. Once SNPs list are created, we can export this list to PLINK statistical tool in the standard format (PED and MAP) files.

gPlink:

gPLINK has integration with PLINK which is a statistical tool for whole genome association analysis. gPLINK is a java based program and is a graphical user interface tool that uses most of the PLINK command operations. It is free software and can be downloaded in this website: <http://pngu.mgh.harvard.edu/purcell/plink/>. The output files of genotyping console (PED and MAP) are used as input files for gPLINK.

A PED file is a white-space (space or tab) delimited file where each line represents one individual and the first six columns are mandatory and in the order 'Family ID', 'Individual ID', 'Paternal ID', 'Maternal ID', 'Sex (1 = male, 2 = female, 0 = missing)' and 'Phenotype (1 = unaffected, 2 = affected, 0 = missing)'. The subsequent columns denote genotypes which can be any character (e.g. 1,2,3,4 or A, C, G, T). 0 denotes a missing genotype. Each SNP must have two alleles (i.e. both alleles are either present or absent). The order of SNPs in the PED file is given in the MAP file, where each line denotes a single marker and the four white-space separated columns are 'Chromosome (1-22, X, Y or 0 for unplaced)', 'Marker name (typically a "rs_" number)', 'Genetic distance in Morgans (this can be fixed to 0)' and 'Base-pair position (bp units)' [73]. These two files are converted to binary PED files (BED) to save space and time. This BED file has FAM file that store the pedigree/phenotype information and create an extended MAP file (.bim) which contains information about the allele names. So this creates (by default): plink.bed (binary file, genotype

information), plink.fam (first six columns of mydata.ped), plink.bim
(extended MAP file: two extra columns = allele names) [74].

Sample based QC: The first step is to load the Bed files of both benign and melanoma samples in gPLINK and then filter them by excluding samples with too much missing genotype data which can lead to poor analysis. The default threshold value of 0.1(10%) is used which means the samples will be excluded if they have more than 10% missing genotype data.

SNP based QC: After sample QC we apply SNP-based QC for each SNP within each sample. First, SNPs with a rate of missing genotype data ≥ 0.05 (5%) are excluded from further analysis. This is to remove any SNP with more than 5% missing genotype data, because if they don't have genotyping data it can lead to false negative, false positive results. Next SNPs with a minor allele frequency ≤ 0.05 are similarly excluded. This is a frequency at which a least minor allele occurs in a population. Any SNP with less than 5% is removed, this filter is applied because SNPs with MAF greater than 5% are said to be involved with diseases. Finally, we exclude SNPs with Hardy Weinberg (HW) deviations observed from expected p-values of 0.01 or greater. HW equilibrium states that allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences. Removing SNPs with extreme deviation from equilibrium. All these QC metrics were performed with recommended threshold values and the SNPs which did not pass these values were removed.

Partek Genomics Suite (PGS):

PGS is a software which is a user-interface, easy to use, has built-in workflows for a variety of genomic workflows that supports Next generation Sequencing, Microarray data and qPCR platforms. We can download the software at their website <http://www.partek.com/updates> , but need license to access it. For studying SNPs, we used Association workflow in the PGS. This workflow requires genotyping calls such as .CHP files from Affymetrix SNP 6.0 Protocol as input for analysis. We performed Sample and SNP based QC for our data and then association analysis for our samples.

The steps were followed as PGS User Association Workflow guide suggested:

The Sample QC option will invoke the sample QC spreadsheet, which shows one sample per row.

1. The rate of missing genotype calls for each sample is given in the Sample NC Rate column. The rate is determined by dividing the number of no calls (NC) by the total number of genotypes in the sample and an unusually high number in this column indicates that the overall genotyping quality in the sample is poor. As a rough guide, one can tolerate a NC rate of up to 5%.
2. The Sample Heterozygosity Rate can also be used as a quality indicator. As a rule of a thumb, you might want to reconsider the samples with the heterozygosity rate which falls out of the interval $\text{mean} \pm 3 \times \text{standard deviations}$ of all the samples. To calculate the mean and standard deviation for the heterozygosity rate of the samples

please use Stat > Descriptive > Column Statistics... and select mean and standard deviation from the list of available Candidate Measures.

3. After removing the samples which did not pass the sample QC criteria you can proceed to the next QC step, Hardy-Weinberg equilibrium, which is essentially QA/QC on SNP level. The resulting spreadsheet (HWE) features one SNP per row. The difference between the observed and expected frequencies of each allele at each locus (or SNP) are tested by χ^2 test (Chi^2 and p-value (Chi^2)).
4. Frequencies of both alleles are provided in the columns A Freq and B Freq, while the minor allele frequency (MAF) corresponds to the one with lower frequency. The remaining three columns contain the no-call frequency (NC Freq), heterozygous frequency (Het Freq), and homozygous frequency (Homoz Freq) at the given locus.
5. Depending on the annotation provided by the array vendor, it may be possible to annotate the SNPs with exact base calls at each locus. Please right-click on a column header and select Insert Annotation. In the Add Rows/Columns to Spreadsheet dialog, please tick mark the Allele A and Allele B boxes. Two new columns will be added to the HWE spreadsheet and will contain the genotype of each allele.
6. At this stage, the following two filters may be considered: 1. A SNP no-call rate should be less than 5%. 2. Minor allele frequency of a SNP should be greater than 5%.

7. After that, have removed the SNPs that are not in Hardy Weinberg equilibrium. For that purpose, a multiple testing correction should be applied to the exact p-value: the cut-off p-value after the correction equals $0.05 / (\text{number of SNPs left after 1st and 2nd filter})$ (in the other words, Bonferroni's correction). The filtering can be performed by the interactive filter (the icon), to first filter in the SNPs with the NC frequency less than 0.05, and then to filter in the SNPs with MAF greater than 0.05. Please note that the effects of the interactive filter are additive.
8. However, in order to proceed to the next step of the workflow, the changes (i.e. filtering in of SNPs which met the chosen QA/QC criteria) have to be applied to the parent spreadsheet; the same SNPs need to be filtered in.
9. To do that please select the parent spreadsheet (in this example this is the one with 17 samples on rows) and then choose Filter > Filter Columns > Filter Columns Based on a List... In the Filter Columns on Spreadsheet dialog, please set the Filter based on spreadsheet to the spreadsheet containing the final SNPs and set the Key column to SNP.

Association Analysis:

After filtering samples and SNPs, now we perform association analysis using chi-square test. The steps are as follows:

10. To set the phenotype (categorical variable) this will be tested for association with the SNPs. In the other words, the allele/genotype

frequencies as specified by the model (please see the discussion below) will be compared between the categories.

11. By setting the Column variable to Tissue, in this example, one will test the association of the SNPs with cancer.
12. The model section allows for specification of the statistical model. Allele: frequencies of alleles (A vs. B) are compared across the categories of the selected variable (i.e. phenotype).
13. Significant p-value indicates that the allele/genotype frequencies are different between the categories of the selected variable, i.e. that an association exists between the genotype and the phenotype.
14. In the present example χ^2 statistic was used to assess the difference in allele frequencies (allele model) between the normal and cancer samples. The resulting spreadsheet (Chi-square) shows one SNP per row.
15. PGS provides the value the associated p-value (chisq p-value) for each SNP.

This p-value along with the genes associated with the filtered SNPs are saved as tab limited text and is exported to IPA.

Ingenuity Pathway Analysis (IPA) is a web-based functional analysis tool for comprehensive genomic data. Quickly gain knowledge on the genes by browsing categorized and curated publication. To download this software go to their website <http://www.ingenuity.com/products/ipa>, which is available for 7 day free trail and have to sign up before you want to download. An email will be sent to your email address which has the link to download the IPA software. Download and install IPA. First import the genes list from PGS.

1. When you hit “run the core analyses” it will show up all the diseases that the genes associated with them.
2. Click on the “cancer” disease and then once a new tab opens where it shows all the cancers, click on the “melanoma” cancer.
3. List of genes and their associated SNPs will show up based on the literature and databases integrated with IPA.

These genes and SNPs can be checked with the list obtained in PGS

III. RESULTS AND DISCUSSION

A total of 27 dermal nevi specimens, preserved in formalin and embedded with paraffin (FFPE), were selected from DLCS. These FFPE block tissues can be stored for a very long time and their accreditation requirements can be fulfilled easily, because of this we were interested in FFPE tissue blocks. The College of American Pathologists (CAP) Commission on Laboratory Accreditation at the present time needs FFPE blocks to be retained for 10 years [75, 76] and also the surgical pathology records. These FFPE tissue blocks can be used for research purposes as long as 1) HIPAA requirements for patient privacy are followed 2) the diagnostic laboratory keeps enough tissue for diagnostic purposes, 3) the diagnostic laboratory has facilities to return any FFPE material remaining after use in research and 4) the demands of applicable institutional review boards (IRB), state and local laws are adhered to. Additionally, CAP currently suggests retaining the archived materials beyond 10 years if possible mainly because of the demand for these specimens for several research areas including cancer research as biomarkers are developing. The clinical data of these specimens, such as patient age, diagnosis and the age of these blocks and their storage at the room temperature for an average period of 8.67 years and average age was 38 years are shown in Table 1. There was enough material found on these specimens for the three extraction methods. For DNA extraction, the sections were cut sufficiently from each sample according to each method (see *Materials and Methods*). The first and last slides were stained with

hematoxylin and eosin (H&E) and observed under microscope (Figure 1) and used for further analysis only if the slides from beginning to the end that were cut have at least 50% of tumors present in each section of sample. Sometimes when we cut we exhaust the tumor tissue in the middle and when we get to the last section we may not have tumor.

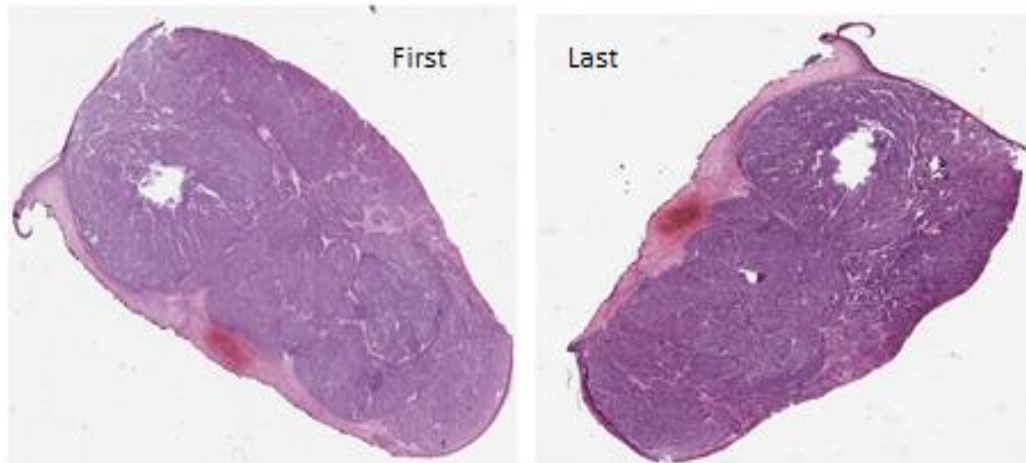


Figure 1: H&E staining for the first and last slides of tissues

Examining the H&E sections on either side of our analyzed sections tells us whether the samples we take for further analysis contain sufficient tumor cells, and allows us to prevent complete depletion of any specimen.

Sample number	Patient age	Sex	Specimen age (years)	Clinical notes
1	30	F	8	Biopsy, nose, left side dermal nevus
2	25	M	8	Biopsy, scalp, right post dermal nevus
3	16	M	8	Biopsy, abdomen left dermal nevus
4	35	F	8	Biopsy, back, left center dermal nevus
5	38	M	8	Biopsy, scalp dermal nevus

6	21	M	8	Biopsy, back right middle dermal nevus
7	26	M	8	Biopsy, abdomen dermal nevus
8	32	F	8	Biopsy, cheek, right dermal nevus
9	31	M	8	Biopsy, axilla, right dermal nevus
10	69	F	8	Biopsy, back, right upper dermal nevus
11	40	F	8	Biopsy, axillary area. Anterior nevus lipomatosus superficialis
12	44	F	8	Biopsy, calf, left post dermal nevus
13	22	F	8	Biopsy, chest, left lateral neurotized dermal nevus
14	56	M	8	Biopsy, cheek, left intradermal melanocytic nevus
15	47	F	8	Biopsy, axilla, left dermal nevus
16	35	F	8	Biopsy, chest dermal nevus
17	29	F	8	Biopsy, back, midline lower neurotized dermal nevus
18	34	F	8	Biopsy, deltoid left ant dermal nevus
19	44	F	8	Biopsy, lip, left upper dermal nevus
20	75	F	8	Biopsy, knee, left medial dermal nevus
21	36	M	8	Biopsy, axilla right dermal nevus
22	60	F	11	Biopsy, back dermal nevus
23	16	F	11	Biopsy, back, left upper dermal nevus
24	41	F	11	Biopsy, forehead, right neurotized dermal nevus
25	16	F	11	Biopsy, back, inferior lower dermal nevus
26	35	M	11	Excision, malar, left dermal nevus
27	74	F	11	Biopsy, back, left upper dermal nevus
Average:	38.04		8.67	

Table 1: Details of the patients used in this study

Aim 1:

In spite of the accessibility of archived specimens, the procedure of fixation leads to the difficulty of using the nucleic acids from these FFPE tissue blocks [77]. So as to overcome these hurdles, it is essential to use effective extraction methods for the nucleic acids. For this, we extracted DNA and compared the three methods from the same 27 FFPE tissues by looking into the quantification measures like total yield, A^{260}/A^{230} and A^{260}/A^{280} ratios, and the purity of DNA. This can be analyzed by looking into the

“amplifiability” of each DNA sample by performing RAPD-PCR to see if the contaminants such as xylene, ethanol and salts are inhibiting the PCR reaction. We found that the quality and quantity of extracted DNA can remarkably affect the downstream processes.

Using the three extraction methods phenol-chloroform isoamyl alcohol, Qiagen-commercial kit and AFA (Adaptive Focused Acoustic), 27 samples were extracted. So many literature suggested different deparaffinization timings and protein digestion times [59, 78-81]. Increasing protein digestion time from overnight to 72 hours can increase the DNA yield according to some studies[77, 79] ,so in the Qiagen QIAamp DNA FFPE Tissue Kit method, the samples were digested with proteinase K for 72 hours instead of overnight digestion in order to get better yield of the DNA. In phenol chloroform isoamyl alcohol method, protein digestion was performed for 72 hours to increase the yield. After protein digestion, the samples were incubated at 94°C for about 40 minutes to reverse crosslinks. The higher the temperature and more incubation time will be better for reversible cross-linking[82],but if the DNA is not cross-linked can cause fragmentation, affect downstream applications [83-85] and also to melt any paraffin that is left in the sample [86]. In order to precipitate the DNA in the samples, they were either kept at -80°C for about 1 hour or can keep at -20°C for about an overnight[59] ,depending on whether you want to proceed the extraction method immediately or later. The glycogen used in this method is to increase the size of the pellets in our samples, but it is avoided since our samples had enough pellet size. Xylene is used in both phenol chloroform and column based extraction methods to remove paraffin. Even though xylene can remove paraffin efficiently[87], any residuals of this can cause low 260/230 ratios and affect

downstream processes like PCR[88]. So it is better to avoid xylene in any extraction process. On the other hand, the AFA does not need xylene or any other organic solvents to remove paraffin. But the deparaffinization takes place in the AFA machine by creating focused bursts of ultra-sonic acoustic energy; this causes hydrodynamic shear stress, emulsification of paraffin and tissue rehydration at the same time. Because of no xylene and automatic removal of paraffin by acoustic energy in this method, the amplicon length in the RAPD PCR (Figure 2 and Table 2) is increased for the AFA samples when compared to other two methods.

<i>Measures</i>	<i>Phenol</i>	<i>Column</i>	<i>AFA</i>
DNA (ng)/section (Qubit)	106.2 ± 33.1	264.3 ± 35.4	134.6 ± 18.1
DNA (ng)/section (Nanodrop)	702.1± 200.2	716 ± 116.1	453.8 ± 53.9
A260/A280	1.94 ± 0.02	2.04 ± 0.03	1.90 ± 0.02
A260/A230	1.71 ± 0.18	1.71 ± 0.12	1.75 ± 0.43
Max. amplicon (bp)	346.7 ± 24.1	347.4 ± 21.4	401.9 ± 10.2
PCR Failure Rate	25.93%	22.22%	3.70%

Table 2: Summary of the three Methods Comparison for 27 samples showing the QC measures

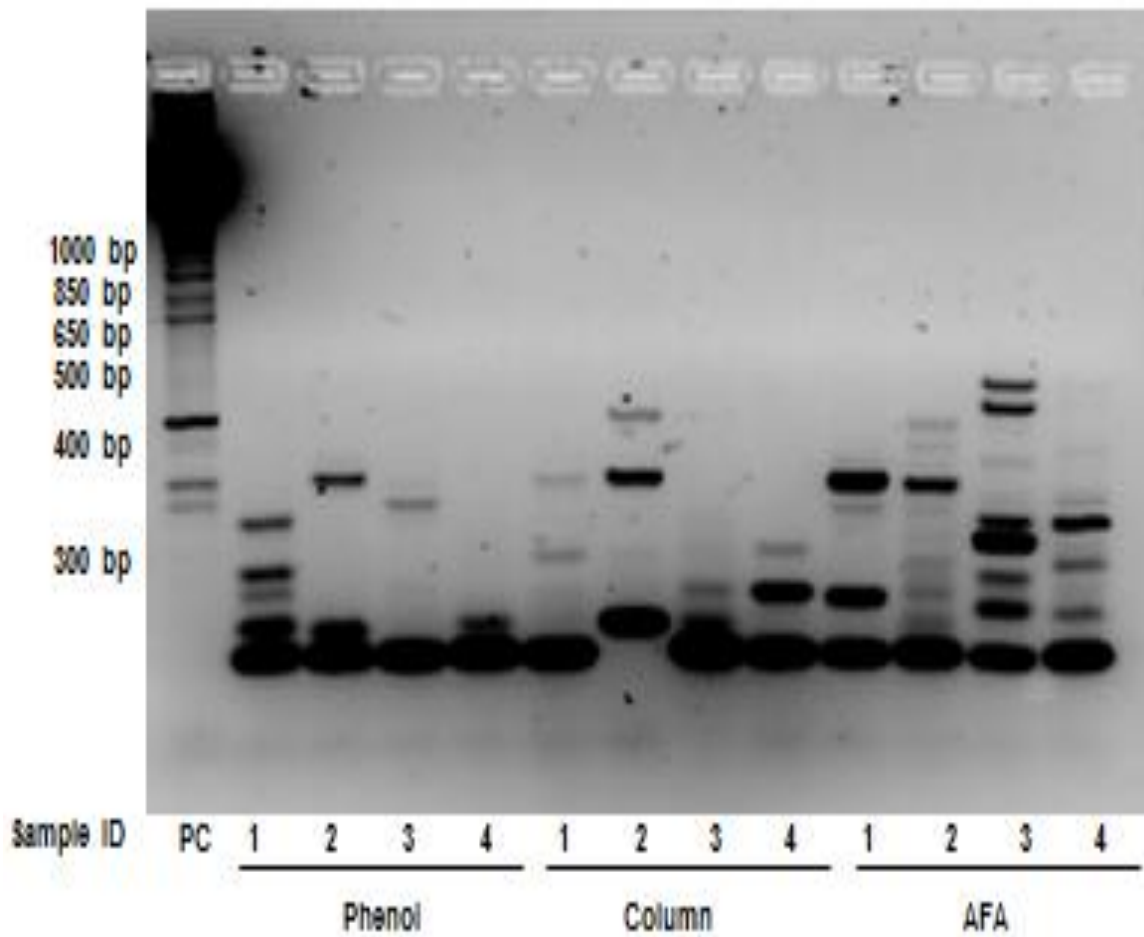


Figure 2: RAPD PCR comparison for the three methods. All the four samples (1, 2, 3, and 4) were compared with negative and positive control (PC; Jurkat DNA) and these same samples were extracted using three different methods

The deparaffinization time in AFA method was increased from 300 sec to 500 sec because initially the paraffin was not removed properly with 300 sec time and so to

improve this step we increased the time. The main factors of these three extraction methods are compared and given in Table 3.

Factors	AFA	Qiagen	PCI
Number of sections (10µm thickness)	~10	~24	~24
Protein Digestion time	~2 hours	~ 72 hours	~ 72 hours
Chemicals	No Organic solvents	Organic solvents	Organic solvents

Table 3: Comparison of factors for the three methods

After extracting the genomic DNA of 27 samples with three methods, the best way to qualify the genomic DNA is to look at the quantification and assessment of dsDNA [89, 90] for downstream analysis. This qualification can be done by checking the contaminants (ethanol, phenol) in the samples with Nanodrop by looking into 260/280, 260/230 ratios. Then look into Qubit concentrations for more accurate concentration of DNA. This is because the dye in the Qubit assay only binds to the double stranded DNA whereas Nanodrop shows high concentration, since it measures absorption at different wavelengths like 230nm, 260nm. For the nucleic acids the absorption takes at 260nm, and so it measures absorption to oligos, single nucleotides, ssRNA rather than just to dsDNA. The 260/280, 260/230, Nanodrop and Qubit concentrations are shown in Table 2. The DNA yield per section for Qubit was also shown in Table 2, Figure 3 and for Nanodrop is shown in Table 2, Figure 4. This is calculated based on the elution volume and the number of sections used per sample.

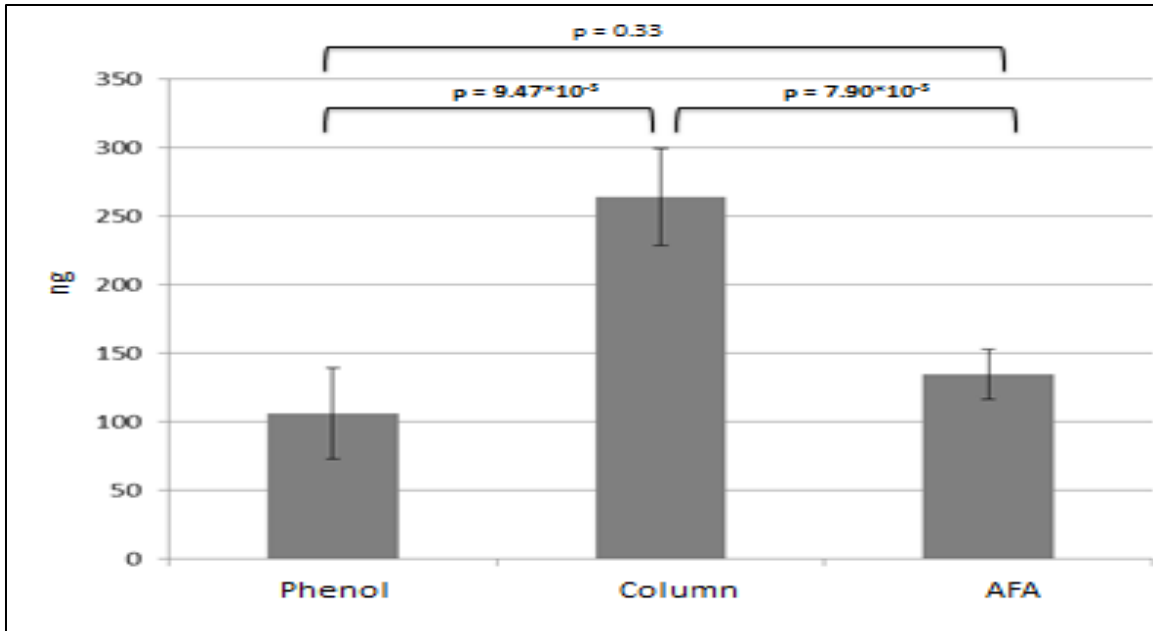


Figure 3: DNA yield per section using Qubit

The column samples showed better yield per section ($p < 10^{-4}$). Even though the column sample showed better yield per section, it is also important to consider the sample purity, reasonable PCR amplicon sizes and less PCR failure rates for using these samples in SNP genotyping, FISH, comparative genomic hybridization and other techniques. Since DNA purity is important to consider for downstream applications, we looked the 260/280 and 260/230 ratios. The 260/280 ratio (Table 2, Figure 4) is the first measure to look for purity of the nucleic acid and it should have ~1.8 value, all the three methods slightly have around 1.8. 260/230 (Table 2, Figure 5) is the second measure for the purity of the nucleic acid and anything around 2 shows the nucleic acids are pure. All three methods have A260/A230

values of approximately 2.0, which is the optimum value of pure DNA

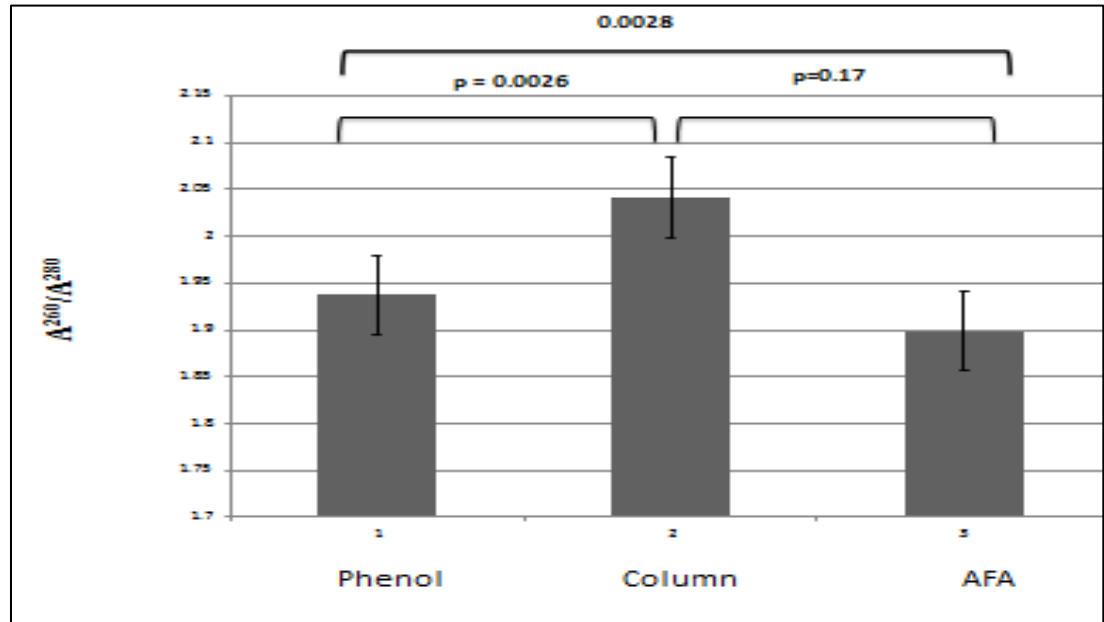


Figure 4: A260/280 ratios comparisons of three methods

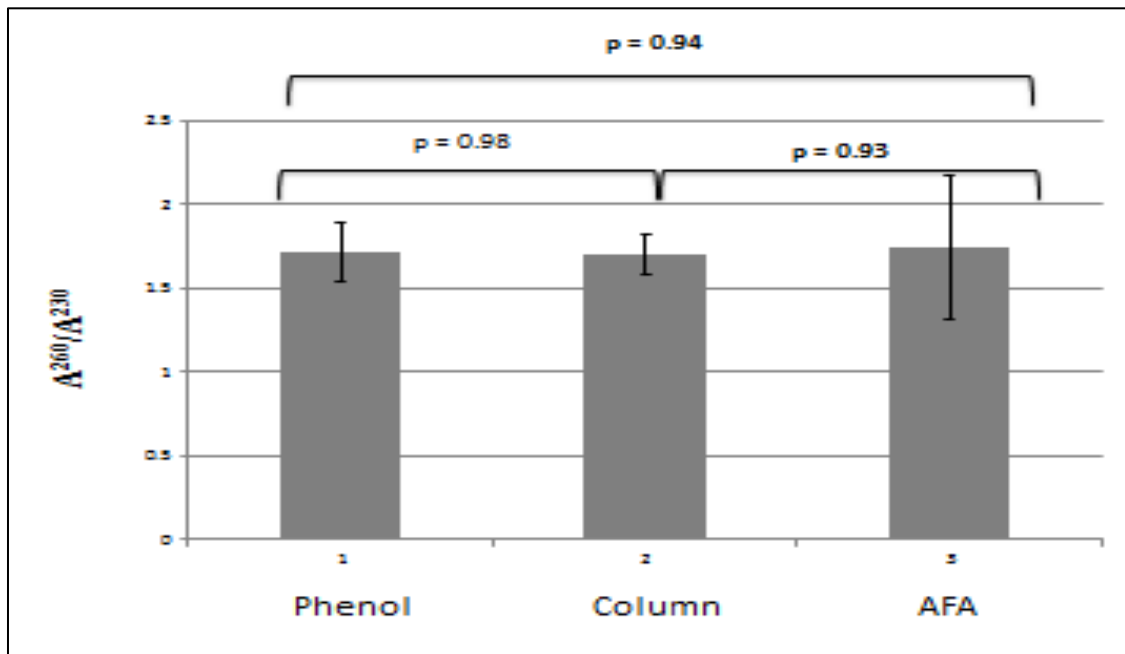


Figure 5: A260/230 ratios comparisons of three methods

To see the size range of the genomic DNA they were separated by 1% agarose gel electrophoresis. For comparison we took the same 4 samples that are extracted with three methods (Figure 6).

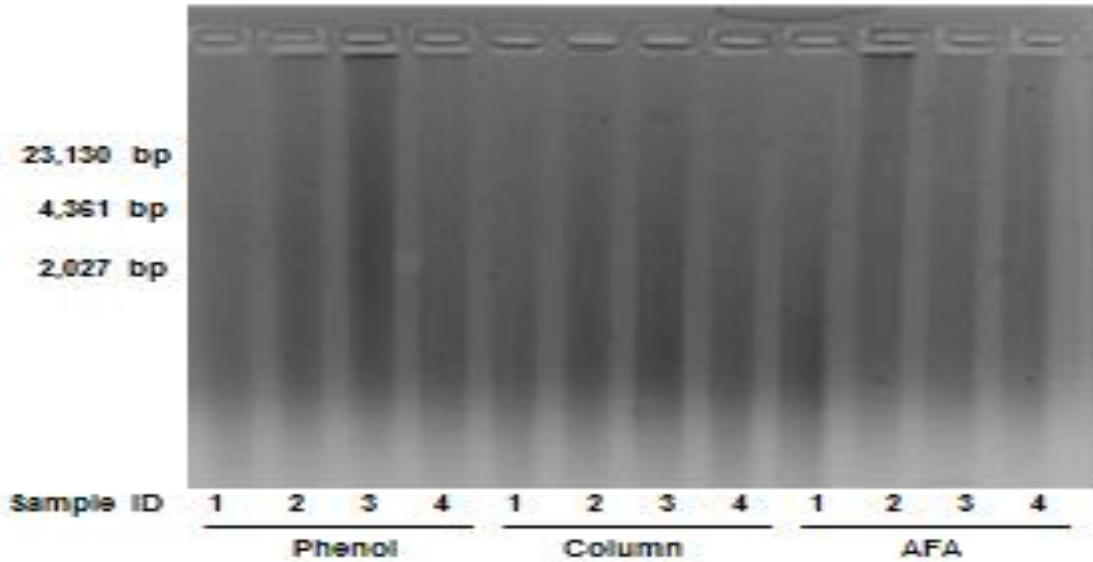


Figure 6: Genomic DNA comparison of three methods

Total amount of the genomic DNA range in the gel is almost similar which shows that the three methods worked the same way. Since RAPD-PCR is one of the good ways to quantify our samples, we now looked at the amplicon size of the 4 samples in Figure and for 27 samples in Table 2. Of all the AFA had showed higher amplicon length compared to other methods(Figure 7, $p \leq 0.04$), this could be because of not using xylene that can interfere with the DNA and can lead to poor PCR analysis and also the PCR failure rates was low (3.70%) for AFA compared to the other two methods (Table 2). Eliminating the contaminants that are carried over during purification process is measured by the rate of “PCR failure” (Table 2).

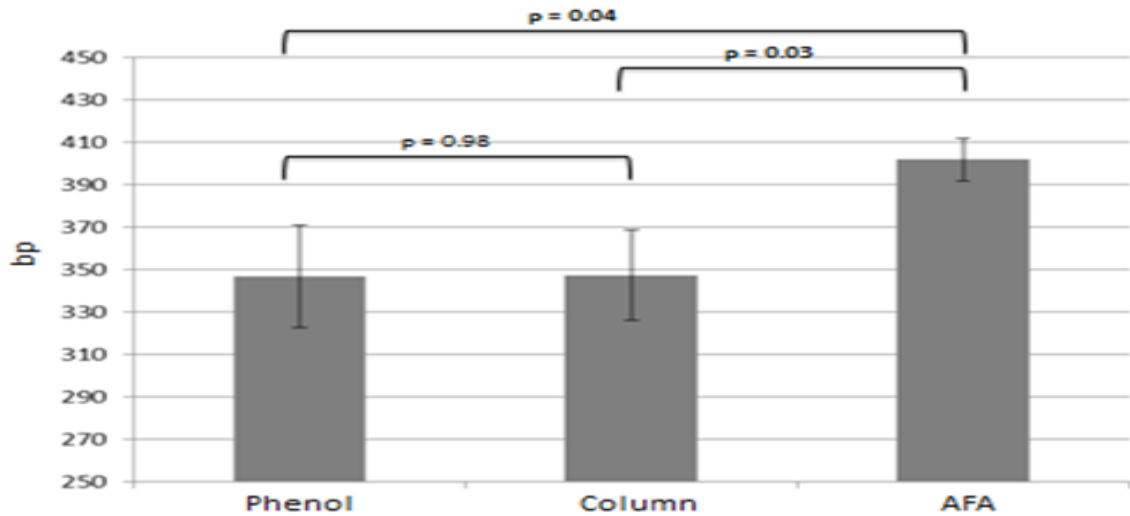


Figure 7: Maximum Amplicon Length in RAPD PCR of three methods

When RAPD-PCR showed amplicons size of less than 300 bp in length, we did not use those samples as it anticipates poor performance in the downstream PCR [91]. The failure rates were around 25% by phenol-chloroform and around 22% by columns. On the other hand, only one of 27 samples (3.7%) failed by AFA.

Aim 2: After extracting good quality of DNA with AFA method from benign FFPE specimens, we used a Qubit fluorometer to determine DNA concentration. Since the concentrations were less in Qubit we used a speed vac to concentrate the DNA. First, by adding distilled water to the each sample to bring up the concentration to as close to 500ng/μl for each Sty and Nsp reactions in digestion and ligation process and one of the advantage of using speed vac is that it will evaporate any leftover solvents like ethanol to avoid carryover contamination. Then the samples were processed and hybridized to the SNP 6.0 array comparative genomic hybridization (aCGH) microarrays. Some studies

have shown that aCGH works well with FFPE specimens [47, 92] for analyzing the tumor tissues [93-95] [96]. Also aCGH is helpful in distinguishing the malignant melanoma from benign cases [45, 97-99] that are ambiguous since distinguishing them can be difficult for the pathologists. Many factors like the age of FFPE tissue block, storage, and also the quality of the DNA obtained from the FFPE tissues can affect the downstream applications like Array CGH. In the SNP 6.0 protocol, we performed QC steps at two stages to see how qualitatively our AFA samples look. One after we perform PCR for these AFA samples and this is shown in Figure 8 Affymetrix suggests to have the PCR products ranged in between 200 and 1100 bp length[72].

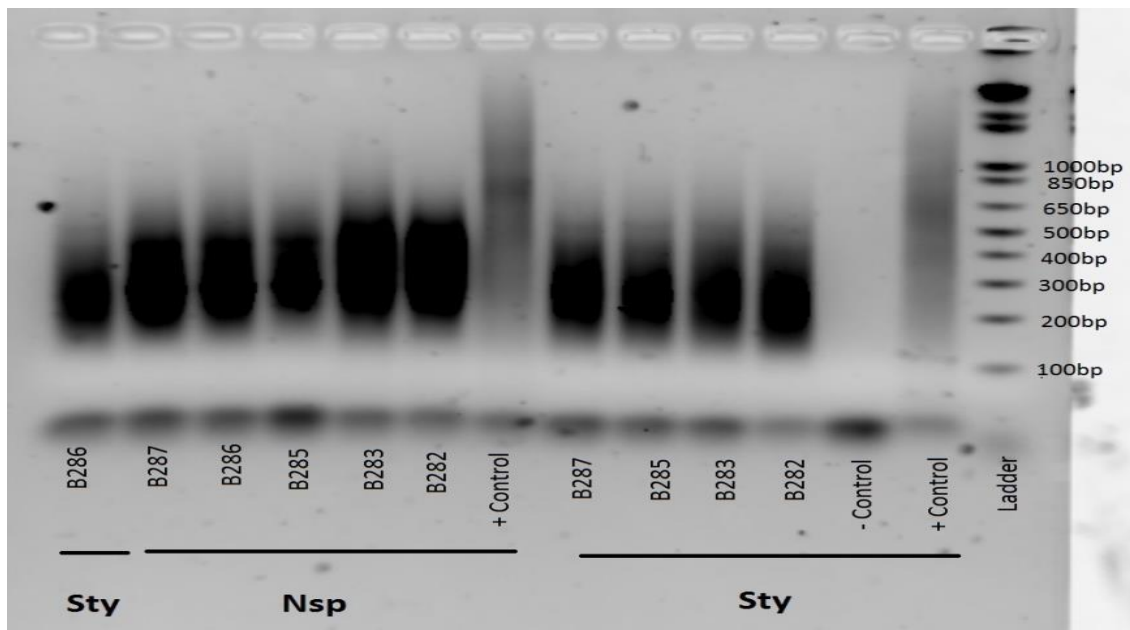


Figure 8: Post PCR gel image in SNP 6.0 Protocol, 5 Sty and Nsp samples were compared with the positive and negative controls

In the 4% agarose gel above the Sty and Nsp samples were in between the range suggested by Affymetrix. So we checked the concentrations and purity of the AFA

samples using Nanodrop .The DNA quantity was sufficient enough for all samples which is around 3500 bp that is suggested by Affymetrix SNP 6.0 protocol and the purity of the samples were also within range (Table 4).

Sample ID	Nanodrop Concentration (ng\μl)	260\280 Ratio	260\230 Ratio
B282	5241.4	1.59	1.87
B283	5380.3	1.47	1.73
B285	3258.1	1.92	2.23
B286	4688.6	1.79	2.09
B287	3962.5	1.87	2.17

Table 4: PCR products of AFA samples showing their Nanodrop concentrations and ratios within the range of SNP 6.0 protocol

After the first QC, we performed the fragmentation step using the PCR to fragment the PCR products so that these small fragments can successfully hybridize to the arrays. Affymetrix SNP 6.0 protocol says that the fragments should be less than 180 bp length to successfully hybridize to the arrays. The samples were run on 4% Agarose gel to assess the fragment size of the samples (Figure 9). All of the samples were fragmented in between 100 and 200 bp.

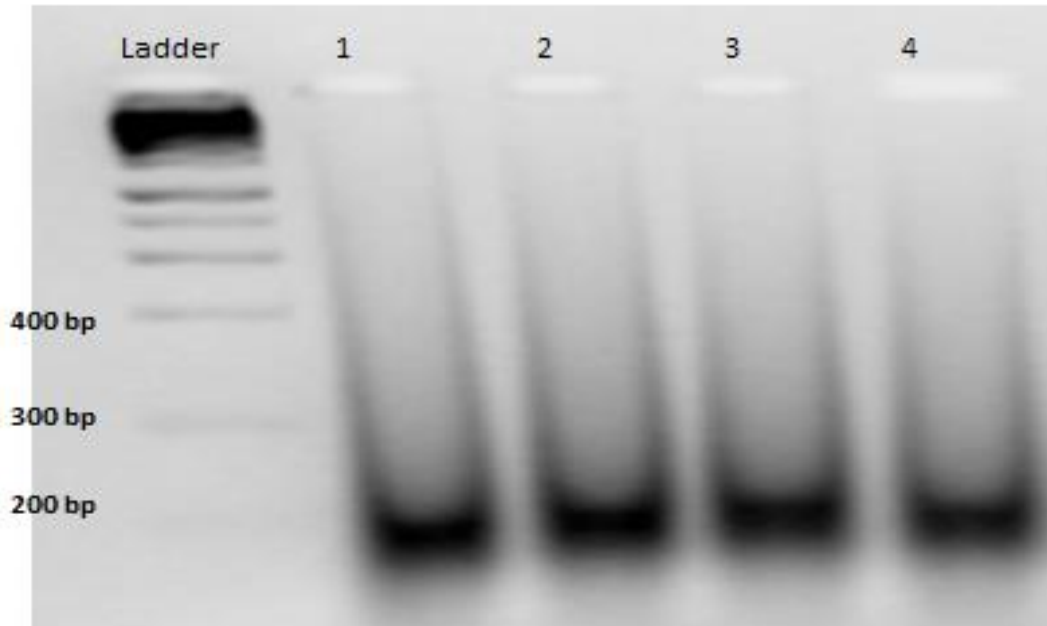


Figure 9: Fragmentation of PCR products on 4% Agarose Gel for 4 samples

(We judged that the quantity of the DNA within the 100-200 bp range was probably sufficient for hybridization, but this could have been further fragmented by repeating fragmentation step for about 30 minutes or could have done this step three times each reaction for 10 minutes rather than one reaction for 30 minutes to increase the DNase I activity [63]. Rather than over-fragment the DNA, it was utilized as it was, and the end results validated this decision.) The fragmented samples were hybridized onto the microarrays, and these microarrays were stained and washed if any fragments that were not hybridized will be washed out. The microarrays were scanned using GeneChip Scanner by Affymetrix and CEL files were obtained.

Microarray Analysis: 9 AFA samples have been able to pass the QC steps in the SNP 6.0 protocol (PCR and fragmentation steps) i.e.; they have showed enough amplicon length, concentrations and fragmented with ≤ 200 bp of length. We used these AFA samples along with the column samples to do next steps. Once CEL files are obtained we further proceeded with microarray analysis by comparing 63 specimens of both columns (54) and AFA (9) samples. The details of these 63 specimens are given in Table 5.

#	Sample	Age at Diagnosis	Year of the sample	Sex	Location	Clark's level	Type
1	M07	77	2008	Male	Right Arm	IV	SUPERFICIAL SPREADING TYPE
2	T10	-	2005	Male	-	-	MALIGNANT MELANOMA
3	T12	80	2005	Male	Forehead	IV	DESMOPLASTIC TYPE
4	T15	75	2005	Female	Upper Back	IV/V	SUPERFICIAL SPREADING TYPE
5	T17	-	2006	Male	-	-	MALIGNANT MELANOMA
6	M18	40	2007	Male	Back	IV	MALIGNANT MELANOMA
7	T19	60	2007	Female	Left Leg	IV	NODULAR SPITZOID TYPE
8	T20	-	2007	Female	Elbow	IV	SUPERFICIAL SPREADING TYPE
9	T23	81	2008	Male	Left Auricular	IV	SUPERFICIAL SPREADING TYPE
10	M27	93	2001	Male	Right Groin	IV	NODULAR MALIGNANT MELANOMA
11	M28	60	2007	Female	Left Leg	IV	NODULAR SPITZOID TYPE
12	M30	-	2011	Male	-	-	MALIGNANT MELANOMA
13	M31	63	2008	Female	Right Forearm	IV	SUPERFICIAL SPREADING TYPE
14	M33	76	2008	Female	Right Cheek	IV	SUPERFICIAL SPREADING TYPE
15	M50	57	2010	Male	Left Arm	IV	SUPERFICIAL SPREADING TYPE
16	M54	45	2012	Female	Left Back	IV	ULCERATED NODULAR MALIGNANT MELANOMA
17	M56	60	2012	Male	Left Shin	IV	SPITZOID TYPE
18	M58	66	2011	Male	Right Cheek		POORLY-DIFFERENTIATED CARCINOMA!!!
19	M59	77	2011	Male	Right Elbow	IV	INVASIVE MALIGNANT MELANOMA
20	M64	27	2011	Female	Right Neck	IV	POSTERIOR INVASIVE POLYPOID MALIGNANT MELANOMA

21	M74	69	2010	Male	Left Temple	IV	SUPERFICIAL SPREADING TYPE
22	M75	74	2010	Male	Left Neck	IV	SUPERFICIAL SPREADING TYPE
23	M76	67	2010	Female	Right Back	IV	Malignant Melanoma with vertical Growth
24	M77	53	2010	Male	Right Arm	IV	NODULAR MALIGNANT MELANOMA
25	M84	-	-	Female	-	-	MALIGNANT MELANOMA
26	M86	54	2009	Female	Anterior Thigh	IV	SUPERFICIAL SPREADING TYPE
27	M98	72	2005	Male	Right Shoulder	III	SUPERFICIAL SPREADING TYPE
28	M107	80	2005	Male	Forehead	IV	DESMOPLASTIC TYPE
29	M111	48	2005	Female	Right Forearm	V	NODULAR VARIANT
30	M113	40	2007	Male	Back	IV	Malignant melanoma
31	M116	74	2007	Female	Right Arm	II	SUPERFICIAL SPREADING TYPE
32	M130	63	2001	Male	Back	IV	SUPERFICIAL SPREADING TYPE
33	M131	45	2001	Female	Left Arm	III	SUPERFICIAL SPREADING TYPE
34	M132	75	2001	Male	Right Chest	III-IV	DESMOPLASTIC TYPE
35	M136	40	2001	Female	Upper Back	III-IV	NODULAR TYPE
36	M137	85	2001	Female	Right Cheek	IV/V	Malignant melanoma
37	M139	87	2001	Male	Right Cheek	V	CONSISTENT WITH THE LENTIGO MALIGNA MELANOMA SUBTYPE
38	M147	90	2002	Male	Left Back	IV	SUPERFICIAL SPREADING TYPE
39	M148	55	2002	Female	Middle Back	II/III	SUPERFICIAL SPREADING TYPE
40	M173	28	2008	Female	Left Shoulder	III	INVASIVE SPITZOID MALIGNANT MELANOMA
41	M174	42	2008	Male	Right Pretibial	III	SUPERFICIAL SPREADING TYPE
42	M191	75	2001	Male	Left Back	IV	MELANOMA WITH REGRESSION
43	NT06	34	2010	Male	Right Arm	-	DERMAL NEVUS
44	NT07	44	2010	Female	Right Buttock	-	DERMAL NEVUS
45	NT12	-	-	Male	-	-	DERMAL NEVUS
46	B15	19	2010	Female	Left Back	-	DERMAL NEVUS
47	B27	39	2011	Male	Left Neck	-	DERMAL NEVUS
48	B29	35	2011	Male	Left Flank	-	DERMAL NEVUS
49	B30	34	2011	Female	Right Neck	-	DERMAL NEVUS
50	B31	44	2011	Male	Right Groin	-	DERMAL NEVUS
51	B47	27	2011	Female	Right Back	-	DERMAL NEVUS
52	B51	62	2011	Female	Right	-	DERMAL NEVUS

Eyebrow							
53	B52	48	2011	Female	Left Neck	-	DERMAL NEVUS
54	B53_3	75	2011	Female	Left	-	DERMAL NEVUS
Shoulder							
55	B254	34	2007	Female	Left Deltoid	-	DERMAL NEVUS
56	B267	33	2013	Female	Right Back	-	DERMAL NEVUS
57	B272	38	2013	Female	-	-	DERMAL NEVUS
58	B279	52	2013	Female	Right Back	-	DERMAL NEVUS
59	B282	41	2013	Female	Right Neck	-	DERMAL NEVUS
60	B283	14	2013	Female	Right	-	DERMAL NEVUS
Shoulder							
61	B285	16	2013	Female	-	-	DERMAL NEVUS
62	B286	21	2013	Male	Right Back	-	DERMAL NEVUS
63	B287	60	2013	Female	Right Neck	-	DERMAL NEVUS

Table 5: Details of the 63 specimens

However, we did not use phenol samples even though they have good amplicon sizes in the RAPD-PCR, they did not have sufficient amplicon length in the PCR, also the concentrations were lower than 3000 ng/ μ l in the SNP 6.0 protocol. Using these CEL files of these 63 specimens we did fragment length analysis by using Affymetrix Power Tools (APT). APT analyzes the microarrays and assigns a call rate to every probe depending on how successfully they hybridized to the microarray. Probes are sorted by the size of the restriction fragment on which they occur. If any piece of region within the restriction sites is missing then this region will also be missing in the samples that are hybridized to the arrays. Any fragment of genomic DNA that is hybridized successfully should have at least a contrast QC of 0.4. The Contrast QC captures the ability of an experiment to resolve SNP signals into three genotype clusters. Here APT measures contrast QC and we compared both AFA and column restriction fragment sizes of Nsp1 and Sty1 (Figure 10). Any fragment size showing below 0.4 threshold of CQC shows that the samples have poor quality. AFA showed largest fragment size of 400 bp in Sty and

500 bp in Nsp compared to the column samples. This APT analysis tells about the quality of the DNA, the more range of fragments present per sample the better the data it will have and can increase the yield of downstream analysis. Even though AFA showed better fragment sizes, they still have missing fragments above 500 bp. This could be because of the starting material (FFPE tissue) quality or samples could have not hybridized properly to the arrays.

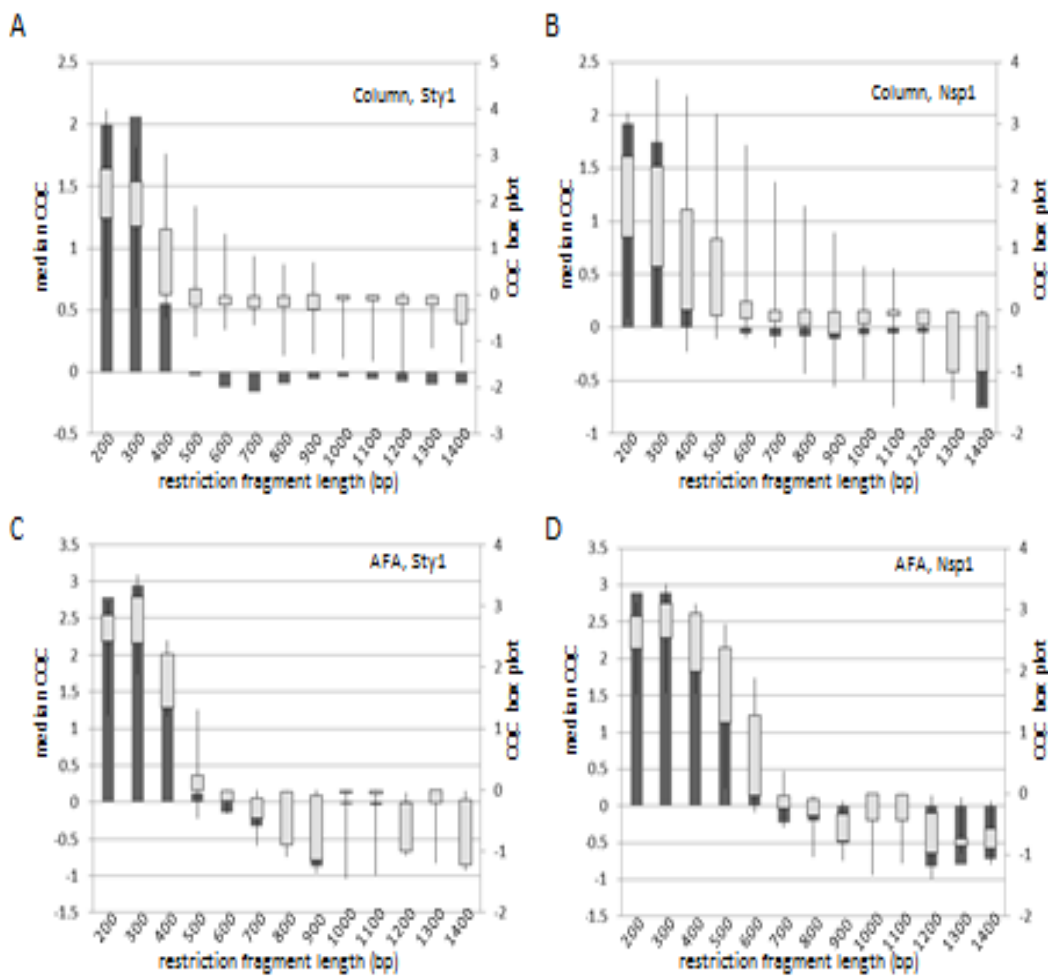


Figure 10: aCGH APT Fragment Length Analysis contrast QC comparison

Aim 3:

All 63 specimens (including benign and melanocytic nevi) were hybridized on to the Affymetrix SNP 6.0 microarrays were analyzed in Genotyping Console (GTC) by Affymetrix, gPLINK, and Partek Genomics Suite (PGS) to perform association analysis.

The 63 CEL files that were obtained after the scanning of microarray chips were opened in Genotyping Console for genotyping analysis. Before performing genotyping, we checked the QC of the samples. The overall quality of samples were tested with contrast QC (CQC) algorithm with a threshold value of ≥ 0.4 . We use this because it is the default CQC value established by Affymetrix for Genome-Wide Human SNP 6.0 array samples.

The Contrast QC captures the ability of an experiment to resolve SNP signals into three genotype clusters. It uses a static set of 10,000 randomly chosen SNP 6.0 SNPs (Figure 11 taken from GTC 4.1 user manual); measuring the difference between peaks in contrast distribution produced by homozygote genotypes, and the valleys they share with the heterozygote peak, and takes the smaller of the two values. In poor quality experiments the homozygote peaks are not well-resolved from the heterozygote peak and the difference values approach zero.

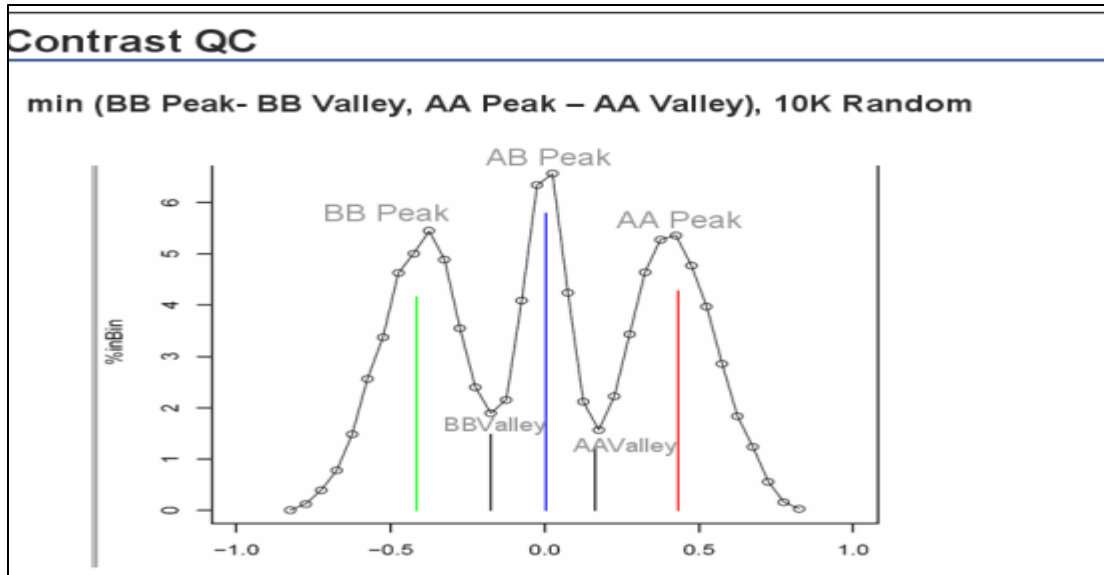


Figure 11: Distribution of Contrast Values. The X axis is the Contrast Value about which a bin of size 0.02 is centered. The Y axis is the % of SNPs (10000 random autosomal GW 6 SNPs) whose Contrast values fall within the bin. Contrast = $\sinh[K*(A-B)/(A+B)]/\sinh(K)$, $K=2$, A and B are the summary values for probes covering the A and B alleles, respectively [1].

Samples passing the contrast QC threshold should show “In bounds” for each sample but all of our samples were “out of bounds” which means they have <0.4 value. Not every SNP is generated by both the StyI and NspI enzymes but some will have the SNPs with just one enzyme set. This issue can be solved by contrast QC. So this Contrast QC values are calculated for contrast distributions produced by a static set of 20K randomly chosen SNPs on NspI fragments and a static set of 20K randomly chosen SNPs on StyI fragments. These are called Contrast QC (NspI) and Contrast QC (StyI), respectively. If the absolute difference between these two values is greater than two, this is evidence that that a sample may have worked properly with one enzyme set, but not with the

other, and the Contrast QC value is adjusted to zero to reflect this problem.

Even though if the samples are “out of bounds” or did not pass the threshold value we can still consider the samples for further analysis (genotyping) if we keep in mind the limitations of the study.

Genotyping analysis identifies genomic variations (SNPs) in the genomes of different individuals. For analyzing SNPs in our samples we performed genotyping analysis. The genotyping results display CHP summary statistics where it shows call rate of all samples. The call rate of a sample is defined as the fraction of the number of SNPs called out of all SNPs on the array. Our samples had an average call rate of 61.73%. The higher the call rate the better we can avoid the false positive/false negatives in the analysis. A good call rate can be greater than 90%. This low call rate could be because of some of the samples did not pass the intensity QC which tells that some samples are missing the expected SNPs on the probes. The genotyping calls of 909,622 SNPs for every sample were determined using the Birdseed V2 genotype calling algorithm, which is in the Affymetrix GTC software. The samples were assigned genotypes and then SNPs were removed when they have low call rate of $\leq 90\%$, any SNP above or equal to 90% means they are good enough for the analysis. The low call rate in our experiment indicates that more than 10% of the SNPs for our samples have missing genotype data. This can cause serious analysis complications downstream. With the help of CHP files, a SNPs list was created for SNPs with call rates $\geq 90\%$ in our samples. The number of SNPs before and after filtering is summarized in Table 5. The flow chart shows the steps in the genotyping console (Figure 10). In order to use these SNPs for further analysis, this list is exported to

gPLINK in the format of PED and MAP files. The PED file has information about the sample like Family ID, Individual ID, phenotype and gender whereas MAP file contains information about marker ID (rs...), chromosome number and base pair position. The quality control (QC) measures are an important step for studying genotype data of genome-wide association studies (GWAS) to reduce false findings.

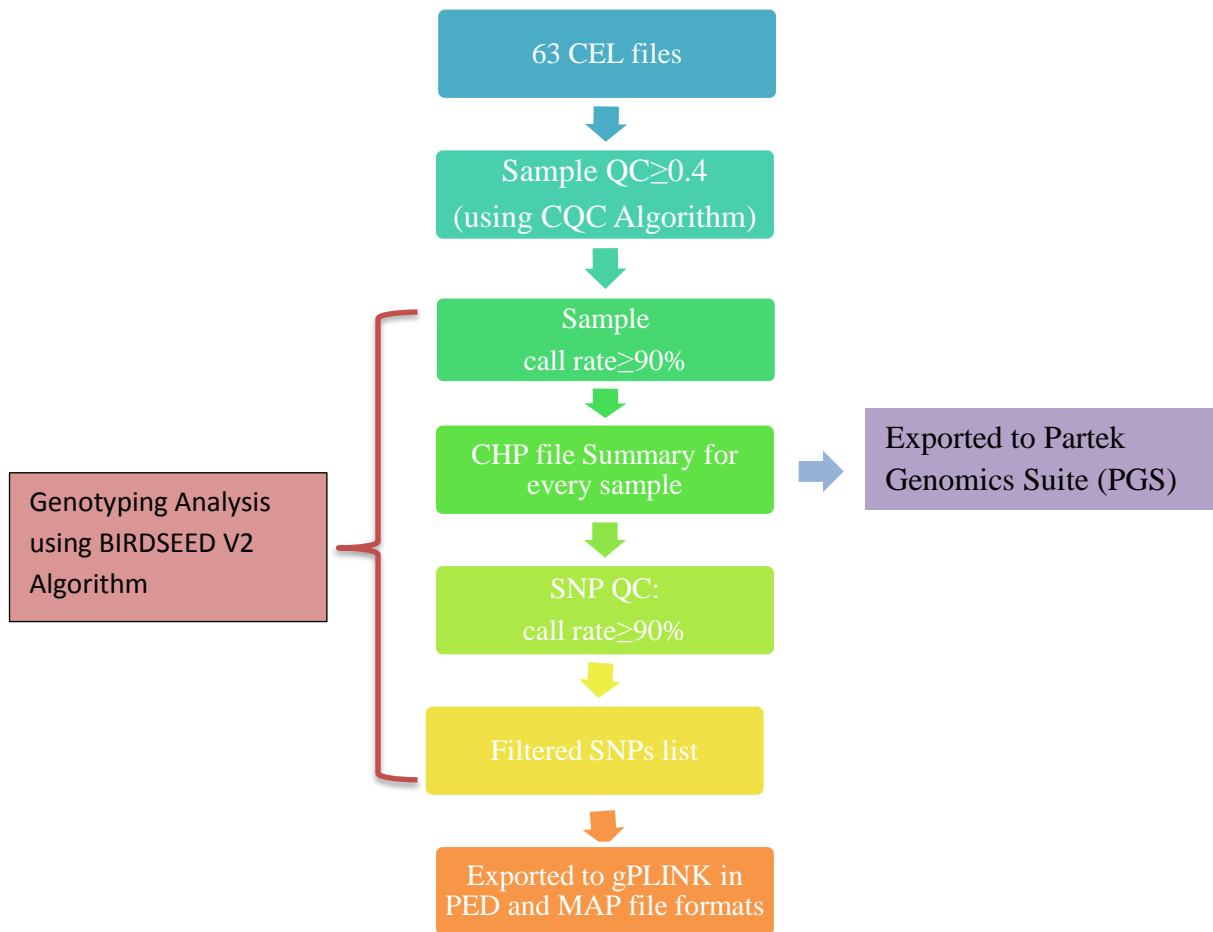


Figure 12: Flow chart showing the QC steps for both Samples and SNPs in GTC

gPLINK

1. The PED and MAP files from GTC were first converted to BED files to bring the information in these two format files together and save space. This BED file is first filtered with recommended QC steps [73, 100]. The recommended SNP QC steps are minor allele frequency, missingness rate, and Hardy Weinberg equilibrium. First, the samples were filtered using missingness rate per individual (MIND) with a threshold of 0.1 this is a good indicator of marker quality. This will remove the samples that have more than 10% missing genotype data since more missing genotyping data leads to false positive or false negative results [100]. But all of our samples have less than 10% missing genotype data. Then we applied SNP based QC for every SNP in every individual. The first SNP based QC step for SNPs is minor allele frequency (MAF). This refers to the frequency at which the least/ less abundant allele occurs in a given population. Any SNP with a 5% or less minor allele frequency are removed. . We use MAF to remove SNPs because the statistical power is very low to detect association for rare SNPs (usually < 1% frequency), so it is better to remove these rare SNPs to avoid burden for the analysis using the power of statistical tools [100] Next SNP based QC is the missing rate per SNP (GENO) with a threshold of 0.05 is used to filter out some more SNPs. If they have more than 5% missing genotype data that can increase the rate of false results. A summary is provided in the Table 6 to show the total number of samples and SNPs excluded and included after QC in the analysis.
2. SNPs were again filtered based on their Hardy Weinberg (HW) equilibrium. HW is a model/law/theorem that states in a population, allele and genotype frequencies remain

constant from generation to generation. The default threshold of 0.05 is used to filter out SNPs.

	Samples in study		SNPs in study	
	Before QC	After QC	Before QC	After QC
GTC	63	63	818245	319777
gPLINK	63	63	316244	107385
PGS	63	63	909622	107909

Table 6: This summarizes the number of SNPs and Samples filtered out after QC steps in each statistical tool

Any SNPs that are deviated from this equilibrium are removed. any deviation from HWE threshold shows that genotyping error happened and/or the controls having association with the diseases [73].Explained with details about the HWE in PGS section.

A difficulty with gPLINK is any sample with “unknown gender” or “wrong gender” will be removed from analysis since gPLINK considers this warning as an error. There are some unknown genders in our samples. Also not every command in PLINK is available in gPLINK like “gender check” to avoid the problem discussed above. PLINK/gPLINK does not have proper contact in case if we want to ask any questions or have any concerns. So as a new starter, it will take time to understand and interpret the results.

Partek Genomics Suite (PGS):

In contrast to gPLINK, PGS does not remove any samples with unknown gender since it takes the CHP summary files from GTC as the input, so we used PGS for further analysis. Partek Genomics Suite is statistical analysis software that is user friendly, fast, memory efficient, can work on large sets of data. It has several built-in workflows for genomic data analysis like microarray and next generation sequencing workflows which are easy to use. For our microarray data (CHP files) we used Microarray Association workflow. Here we did sample and SNP-based QC steps. 63 Samples were filtered using the no call rate (NC) for sample which means the rate of missing genotype calls for every sample; it is a good way of assessing the sample quality. It has a default threshold of $\leq 5\%$, any sample with greater than 5% missing genotype data are removed since it can cause false-positive and false-negative results and sample heterozygosity rate which is the proportion of heterozygous genotypes for an individual [73] is an indicator of sample DNA quality. the threshold for this rate is 0.908412 (max)–0.08313(min). The heterozygosity rate of all samples should fall in the interval of $\text{mean} \pm 3 \times \text{standard deviation}$. Any sample out of this range is removed which means the samples have less DNA quality [73] After this, we performed SNP based QC for every SNP in each sample. Hardy Weinberg equilibrium, a SNP based QC step is a model which states that the genotype and allele frequencies of a given population will be constant from generation to next in the absence of other evolutionary influences. any deviation from HWE threshold shows that genotyping error happened and/or the controls having association with the diseases [73]. Usually HWE is tested first for controls because they will have similar frequencies (allele and genotypes) than cases. The deviation can be determined by p-

value with a default threshold of > 0.05 . p-value is used to test statistical hypothesis of the observed sample results. Any SNP (in controls) with the observed p-value is less than the expected p-value(0.05) is deviated from the HWE and thus excluded from the further analysis [73]. This p-value is compared against the allele/genotype frequencies of controls. Before performing the HWE QC step, we have to filter SNPs with No call NC frequency for SNP which also gives a good assessment for every marker quality, is the rate of missing genotype calls for every SNP; any SNP will be removed if it has greater than 0.05 missing genotype data to avoid spurious results. Then the next QC step is minor allele frequency (MAF). This refers to the frequency at which the least/ less abundant allele occurs in a given population. The default threshold is greater than or equal to 0.05 which means any SNP with less than 0.05 MAF value is removed. We use MAF to remove SNPs because the statistical power is very low to detect association for rare SNPs (usually $< 1\%$ frequency), so it is better to remove these rare SNPs to avoid burden for the analysis using the power of statistical tools [100] Once SNPs are removed with the NC frequency, MAF and p-value of HWE, then we used an allelic chi-square test, this chi square test is used to determine if there is any significant differences between observed and expected allelic frequencies between cases (melanomas) and controls (benign). In this test, the SNPs were filtered again based on chi-square p-value using a cutoff p-value of 0.05. This p-value here is compared against cases and controls. Any SNP that has a p-value greater than 0.05 was excluded from the analysis. Significant p-value (≤ 0.05) shows that the allele frequencies are different between the cases and controls, which mean there is an association exists between the allele or marker and the increased risk of disease. Once we know that the samples (controls) are in HWE then we compare these

controls to our cases. This test is just to see whether the samples have any poor genotyping data, or have any inbreeding which can lead to biased comparison with other population (cases). The parameters and thresholds are recommended by the PGS and shown in Table 7. After these steps, we have the filtered SNPs list along with their associated genes and chi square p-value. All the steps after using GTC are given in Figure 13 until we get significant SNPs.

Association Analysis Tools	Sample QC & Thresholds	SNPs QC & Thresholds
gPLINK	Missingness rate per individual (MIND) $\leq 10\%$	Missingness rate per SNP $\leq 5\%$ Minor allele frequency (MAF) $\geq 5\%$ Hardy Weinberg Equilibrium (HWE) $< 5\%$
PGS	No call Frequency $\leq 5\%$ Sample Heterozygosity rate = 0.908412 (max) – 0.08313 (min)	No call frequency $\leq 5\%$ Minor Allele Frequency (MAF) $\geq 5\%$ Chi-square p-value $\leq 5\%$

Table 7: Statistical Tools with their parameters and thresholds

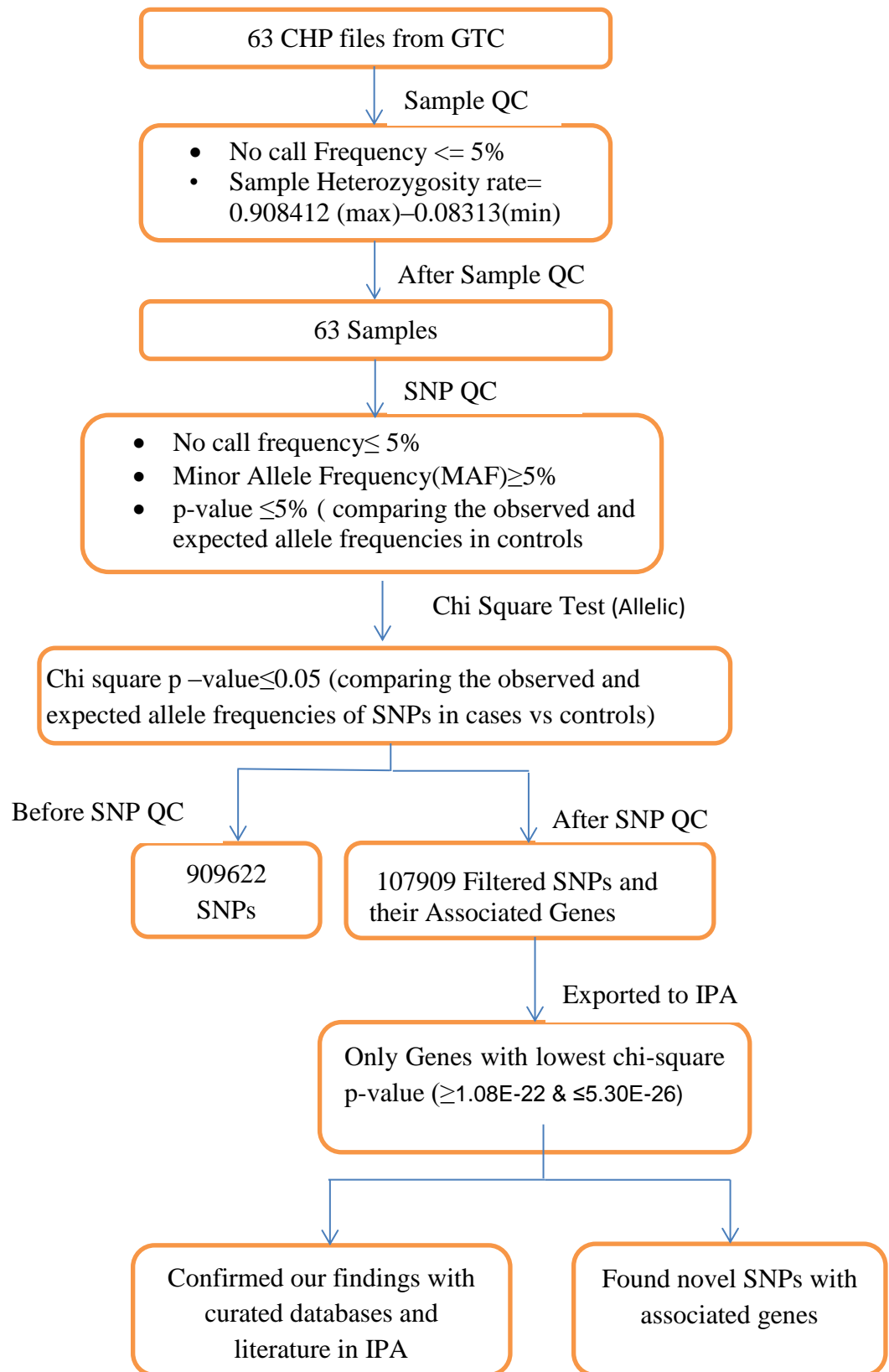


Figure 13: Flowchart showing the steps in getting the filtered SNPs using PGS and confirming them using IPA.

IPA:

Ingenuity Pathway Analysis is a web based functional analysis tool for comprehensive genomic data. We are interested in IPA because we want to explore the significance of our findings. The genes list is being exported to Ingenuity Pathway Analysis (IPA) to look whether the genes are known to be involved in cancer (especially melanoma) by taking the genes with SNPs having the lowest associated p-values. By making the chi-square p-value even stricter - , filtered out the SNPs within the range of $\geq 1.08E-22$ & $\leq 5.30E-26$ and we limited the analysis to 7,389 genes. The SNPs and genes list from chi-square allelic test are shown to be involved in the melanoma in the IPA based on several databases like COSMIC, OMIM, Gene Ontology. These genes along with the SNPs shown in IPA were compared back to the corresponding SNPs list we got from the chi-square test in the PGS to validate our findings. We found some genes and their associated SNPs in the melanoma that are published in the literature are also present in our samples. Some SNPs are found in both melanoma and benign samples and some are just in melanoma. These SNPs in benign may predispose the benign to malignant melanoma according to a literature [54]. These are listed in the Table 8.

SNP ID	GENES	References
rs1267649 (B&M)	BRAF	Meyer, P., C. Sergi, and C. Garbe, <i>Polymorphisms of the BRAF gene predispose males to malignant melanoma</i>
rs401681(M)	CLPTM1L	Stefanaki, I., et al., <i>Replication and predictive value of SNPs associated with melanoma and pigmentation traits in a Southern European case-control study</i>
rs4845618(B&M)	IL6R	Gu, F., et al., <i>Interleukin and interleukin receptor gene polymorphisms and susceptibility to melanoma</i>

rs2736100(M)	TERT	lles, M.M., et al., <i>The effect on melanoma risk of genes previously associated with telomere length</i>
rs9420907(B&M)	OBFC1	lles, M.M., et al., <i>The effect on melanoma risk of genes previously associated with telomere length</i>
rs10757257(B&M)	MTAP	Kvaskoff, M., et al., <i>Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma</i>
rs1015362(B&M)	ASIP	Maccioni, L., et al., <i>Variants at chromosome 20 (ASIP locus) and melanoma risk. Int J Cancer</i>
rs2284063(B&M)	PLA2G6	Kvaskoff, M., et al., <i>Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma</i>

Table 8: Genes and associated SNPs from IPA shown to be involved in melanoma.

Some SNPs are found both in melanoma and benign (B&M) and some just found in melanoma (M). CLPTM1L rs401681 are originally found to be in lung cancer as well as melanoma [101] and TERT rs2736100 [102] both found in melanoma cases only whereas BRAF rs1267649 [103] , IL6R rs4845618 [104], ASIP rs1015362 [105] , OBFC1 rs9420907 [102]. MTAP rs10757257 [54] and PLA2G6 rs2284063 [54] are found in both benign nevi and melanoma.

CLPTM1L rs401681

CLPTM1L gene is associated with the cisplatin-induced apoptosis and it lies in the cancer susceptibility locus on chromosome 5p 15.33, this is usually over expressed in the

melanoma. rs401681 showing a significant p-value of 9.6×10^{-6} is found in this gene and shown to be involved in many cancers like breast cancer [106], colorectal cancer [106] and also melanoma [106].

TERT rs2736100

TERT is a telomerase reverse transcriptase and a ribonucleoprotein polymerase which maintains telomerase ends located in chromosome 5. This telomerase functions as cellular senescence and results in shortening of telomeres. This shortening of telomeres are involved in the risk of cancer [102] especially melanoma [102]. rs2736100 having a p-value of 0.02 and MAF value of 0.486 is located in the TERT gene and recently is found to be involved in glioma [107], thyroid [107], bladder [107] cancer and found in melanoma [107].

BRAF rs1267649

BRAF is a serine/threonine kinase in Ras/Raf/MAPK signal transduction pathway [108, 109] involved in the cell growth and is a proto oncogene. If mutated, the cell growth will be continuous and leads to tumorigenesis. It is located on 7q34 chromosome with 190kb of length. Recent studies have shown that this gene is mutated in the malignant melanoma (66%). Also the BRAF mutation in benign can give rise to malignant melanoma [110]. Meyer et al showed that rs1267649 is having a significant minor allele frequency of ≥ 0.1 and is one of the non-coding SNPs present in intron 5 in the BRAF gene. This SNP is shown to be involved in the increased risk of malignant melanoma in their study.

IL6R rs4845618

IL6R is a tumor growth inhibitor for early-stage of melanoma and a growth factor for the tumor cells for advanced stage [104] patients with malignant melanoma when pretreated with serum IL6 have a very short survival time [104]. It is present in chromosome 1. IL6 shows its activity when bound to IL6R which alters gene expression. Gu et al found that IL6R along with rs4845618 is involved in the melanoma risk with a p-value in Hardy Weinberg equilibrium (HWE) is 1.04-2.84, is present in the intron region of the IL6R.

ASIP rs1015362

ASIP gene has a role in melanogenesis [111], when ASIP binds to MC1R it initiates signaling and blocks the cAMP production. This leads to the down regulation of eumelanogenesis (black/ brown pigments) and eventually increases pheomelanin (yellow/red pigments). It is located in chromosome 20 and it has shown that the rs1015362 has a MAF value of ≥ 0.1 and is located 110kb in the ASIP gene [105]. It is most likely associated with the sun sensitivity like sunburns. So this SNP has an effect on malignant melanoma.

OBFC1 rs9420907

OBFC1 is a subunit of alpha accessory factor that stimulates the activity of DNA polymerase-alpha-primase; this enzyme initiates the DNA replication. Mutations in this gene leads to cancer [112]. It is located in chromosome 10q24.33. This gene is present in the longer telomere length and has shown to be involved in cancers like melanoma as well as nevi and also breast cancer. rs9420907 which is present in the OBFC1 gene is associated with melanoma with a significant p-value of 0.001.

MTAP rs10757257

MTAP is a gene which plays a major role in polyamine metabolism and very important in the salvage of adenine and methionine. The deficiency in this enzyme leads to cancer because the p16 tumor suppressor gene and this MTAP gene are deleted together usually. Located in 9p21 and 22q13 and is shown to be in the benign nevi and has a high risk of melanoma [54] .

PLA2G6 rs2284063

PLA2G6 is a phospholipase A2 gene involved in the Ras signaling pathway. Mutations in this gene or Ras signaling leads to the tumorigenesis. This is shown to be involved in benign and the risk of melanoma [54]. rs2284063 is associated with this gene in benign as we as melanoma [54]

Also in IPA we found some genes that are known to be involved in melanoma, but with SNPs in our list which are different from the published cancer-associated SNPs. These are not known to be involved in melanoma so far. So we expect that these SNPs may be involved in melanoma (Table 9), or may be in linkage disequilibrium with cancer-associated SNPs.

GENES	SNP ID(New)
ERC2	rs732887 (B&M)
LRP1B	rs1372254 (B&M)
PDE1A	rs16823163 (B&M)

SMAD3	rs9492489 (B&M)
CDH13	rs11860430 (M)
GRIN2B	rs10772713 (B&M)
CLCA1	rs1358826 (B&M)
GNAQ	rs17063991(B&M)

Table 9: Genes and new SNPs found in either benign and nevi or just melanoma samples

Table 9 shows *ERC2* [113], *LRP1B*[114], *PDE1* [115], *SMAD3* [116] repression of *CDH13* [117] , *GRIN2B* [118], *CLCA1* [119], *GNAQ* [56]. All these genes are known to be involved in melanoma. It may be that the SNPs associated with them are also involved in melanoma.

ERC2 It is a regulator of neurotransmitter release located in chromosome 3p14.3. Frequent transcriptional and genetic inactivation of *ERC2* can be involved in carcinogenesis [120] and in melanoma it is said to be deleted [113, 121] .rs732887 is found in this gene with significant p-value of showing that it may have some association with the melanoma [113].

LRP1B

This gene belongs to low density lipoprotein receptor gene family, located on 2q22.1 chromosome. They play an important role in the cell functions and development. This gene is involved in the melanoma [114].

PDE1

PDE1 is a phosphodiesterase calmodulin dependent protein, located on chromosome 2q32.1. The role of PDE1 gene is to degrade cGMP and cAMP. This gene regulates the cell proliferation, if over expressed it leads to cancer [122] and melanoma [115].

SMAD3

This gene regulates the cell death, differentiation and proliferation. This helps in the tumor growth in cancer. Repression of this enzyme leads to cancer [123, 124]. It is located in chromosome 15q22.

CDH13

It is a protein coding gene that regulates the cell survival, proliferation and growth and repression of this gene can cause tumor progression[125] and eventually to melanoma [117]. Located on 16q23.3 chromosome. It is indirectly involved in ERK signaling in the melanoma [126]

GRIN2B

GRIN2B is a protein coding gene of glutamate receptor .Located on 12p13.1 chromosome. It has involvement in Ras signaling pathway [127]. GRIN2B is one of the genes that cause melanoma when mutated according to a study [128].

CLCA1

It is a protein coding gene located on 1p22.3. Plays an important role in cell adhesion. Involved in basal cell adhesion and/or in squamous epithelia. It can act as a tumor suppressor in colorectal and breast cancer. It has a key role for cell adhesion in the early stages of lung metastasis [129]. It is a tumor suppressor gene but when mutated leads to melanoma [130].

GNAQ

It is a Guanine nucleotide-binding protein which acts like transducer in various transmembrane signaling pathways. Located on 9q21.2 chromosome. It has involvement in MEK/MAPK and/or PI3K/AKT signaling indirectly [131]. The frequent somatic mutation of this GNAQ cause melanoma [56] . It is a therapeutic biomarker for the uveal melanoma [56].

However the SNPs associated with all these genes were so far not shown in the literature that they involved in cancer/disease. But these SNPs have significant chi-square p-value with in the range of $\geq 1.08E-22$ & $\leq 5.30E-26$, which indicates that they may be involved in the disease. Further studies are needed however.

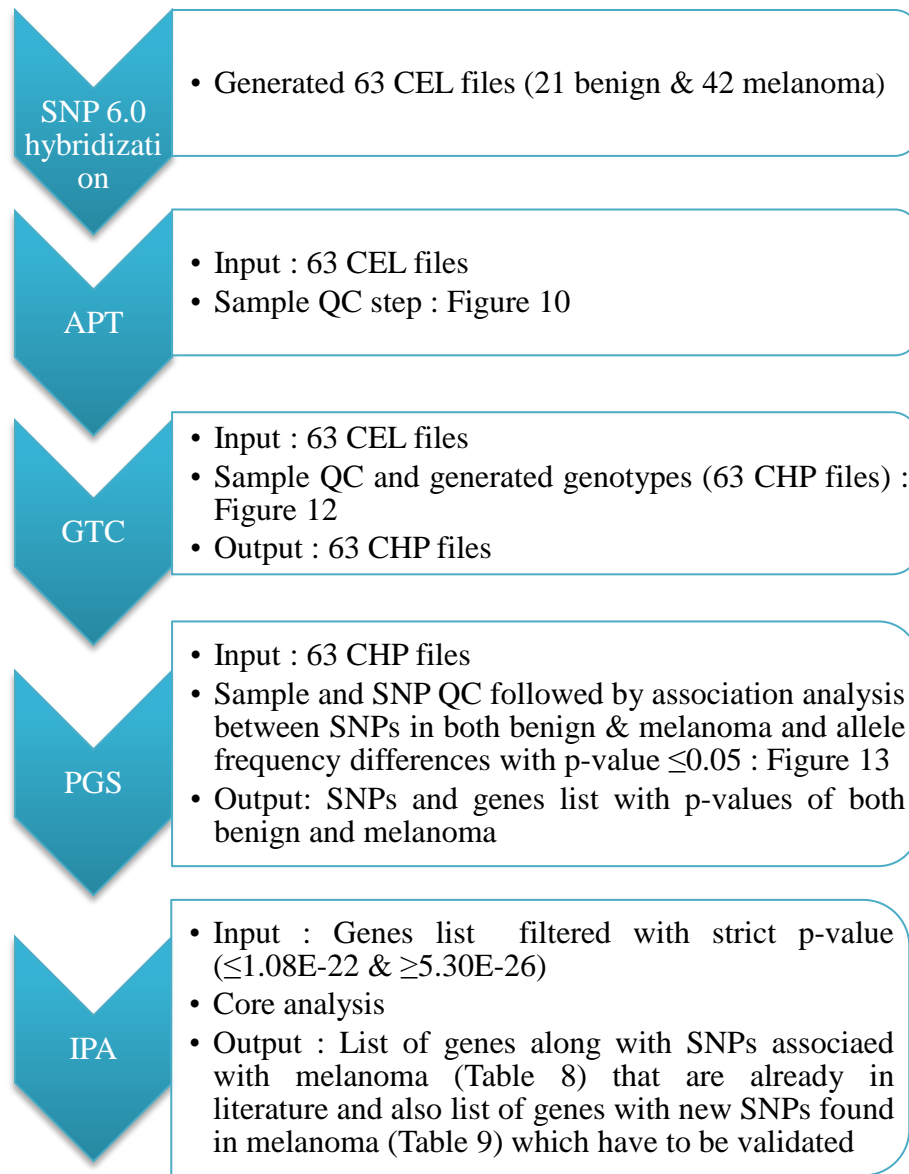


Figure 14: Flowchart showing sequential steps in analysis

Flowchart summarizes the sequential steps in filtering and analyzing the SNPs in both benign and melanoma using different statistical tools.

III. CONCLUSIONS

This part summarizes the findings and the problems we solved in the thesis work and possible future directions.

In this thesis, we showed that the type of DNA extraction method does have an effect on DNA quality, since getting DNA from FFPE tissues is always challenging even though it has several advantages like stable for decades, can use very old FFPE blocks, vast number of blocks available, easily accessible and easy handling and used especially in the diagnosis of cancer. Column is a solid-phase extraction and phenol-chloroform method is liquid-liquid extraction, both these methods require more sections of FFPE tissues, also these methods de-paraffinize the tissues using organic solvents which can later interfere with the DNA purification process and inhibit PCR reaction. On the other hand, AFA which is solid-phase extraction that requires less number of tissue sections compared to other two methods. This AFA method de-paraffinizes the tissues using acoustic energy without any organic solvent xylene, which is commonly used for removing paraffin in other two extraction methods. Our results show that the use of xylene is associated with decreased amplicon length in PCR and increased RAPD PCR failure rate [132]. Any sample shown less than 300 base pair length by RAPD PCR is considered as failure here. It is very important to have greater than 300 base pair length of amplicon for downstream analysis like array CGH [133] especially when using tumor biopsies of limited size.

Although xylene-free extraction methods have shown to be successful, they are not standardized to reproducibility and safety guidelines for using clinically [134, 135].

Also in the AFA method, tissue rehydration takes place during the de-paraffinization process that allows proper protein digestion and complete reverse-crosslinking, supporting efficient purification of DNA [136].

The improved DNA quality from FFPE tissues is very important for the array CGH analysis. One of the studies showed that the quality and quantity of FFPE DNA is assessed by using Nanodrop, 260/230, 260/280 ratios and gel electrophoresis for looking into genomic DNA for hybridizing successfully them to the microarrays. However they did not perform PCR to know the degradation of FFPE DNA samples [137]. Here but in our studies we did perform RAPD PCR along with Nanodrop, qubit and gel electrophoresis to assess the quality of the samples. Of all the three extraction methods, AFA method showed improvements in the FFPE DNA quality. Ours is the first study to compare AFA with column and phenol methods.

AFA samples were performed with SNP 6.0 protocol, successfully hybridized to the microarrays and were compared against the column samples, the phenol samples were not used because they did not pass the initial QC steps in the SNP 6.0 protocol. The AFA samples showed better initial quality control metrics like the optimum amplicon length in the PCR, enough concentrations to proceed to next steps, good fragmentation length so that they can hybridize to the microarrays successfully. These microarrays of AFA and column samples when analyzed with the APT fragment length analysis, AFA samples with fragments are hybridized better than column samples. This proves that extraction

method of FFPE DNA does help in improving the quality of the DNA for hybridizing onto the microarrays. The raw array data that was generated with these AFA samples after SNP 6.0 protocol were genotyped using genotyping console and then the SNPs were filtered and analyzed in Partek Genomics Suite. The SNPs with their genes from Partek were confirmed in the Ingenuity pathway analysis tool. Some of the SNPs along with their genes in our list were already shown in literature in Table 8 that they are involved in the melanoma and some of them were shown in our benign samples. BRAF along with rs1267649 is one of the most common genes and SNPs associated with melanoma is also shown in our list. Also found some new SNPs in our list that were not shown to be involved in the melanoma so far but the genes associated with these SNPs. Some of these SNPs are found both in the melanoma and benign samples and some are just in the melanoma. Benign nevi which are non-cancerous and not metastatic can give rise to malignant melanoma when exposed to certain risk factors [54] whereas malignant melanoma is a metastatic cancer. So these findings suggest that the new SNPs around the known genes may also be involved in the melanoma and also the SNPs that are in the benign nevi may give rise to the malignant melanoma. Recent findings showed that the nevi, which are a risk factor for melanoma, have distinct patterns of genes and SNPs. These findings may help in identifying the subsets of nevi that have distinct biological features and the risk phenotypes in the individuals, however more studies have to be done to confirm their findings [138]. The project future directions include: Increase the size of the samples, perform genotyping analysis separately for benign and melanoma with normal samples. Find whether the SNPs found in our study are driver mutations or not by linkage disequilibrium (LD) plots. Look into literature to find the function and also

whether the genes we found in our data that are involved in melanoma are also involved in other cancers. Perform Next Generation Sequencing (NGS) to generate genotype data. SNPs obtained from clinical specimens that are statistically significant in malignant melanoma and/or benign nevi, may represent the functional genes involved in melanoma. Significant SNPs in benign nevi and/or melanoma that may predispose to malignant melanoma can improve vigilance by encouraging detection of the melanoma in the early stages and prognostic decisions.

III. REFERENCES

1. Affymetrix, *Distribution of Contrast QCs*, www.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf.
2. Olszanski, A.J., *Current and future roles of targeted therapy and immunotherapy in advanced melanoma*. J Manag Care Spec Pharm, 2014. **20**(4): p. 346-56.
3. Dabas, N., et al., *Diagnostic role of chromosomal instability in melanoma*. J Skin Cancer, 2012. **2012**: p. 914267.
4. Lee, J.A., *Melanoma and exposure to sunlight*. Epidemiol Rev, 1982. **4**: p. 110-36.
5. Jensen, O.M. and A.M. Bolander, *Trends in malignant melanoma of the skin*. World Health Stat Q, 1980. **33**(1): p. 2-26.
6. Swerdlow, A.J., et al., *Benign melanocytic naevi as a risk factor for malignant melanoma*. Br Med J (Clin Res Ed), 1986. **292**(6535): p. 1555-9.
7. Schneider, J.S., D.H. Moore, 2nd, and R.W. Sagebiel, *Risk factors for melanoma incidence in prospective follow-up. The importance of atypical (dysplastic) nevi*. Arch Dermatol, 1994. **130**(8): p. 1002-7.
8. Holly, E.A., et al., *Number of melanocytic nevi as a major risk factor for malignant melanoma*. J Am Acad Dermatol, 1987. **17**(3): p. 459-68.
9. Greene, M.H., et al., *High risk of malignant melanoma in melanoma-prone families with dysplastic nevi*. Ann Intern Med, 1985. **102**(4): p. 458-65.
10. Masters, B.R., et al., *Rapid observation of unfixed, unstained human skin biopsy specimens with confocal microscopy and visualization*. J Biomed Opt, 1997. **2**(4): p. 437-45.
11. Institute, N.C., *Cancer Staging*. 2013.
12. Institute, N.C., *Tumor Grade*. 2012.
13. Gerami, P., et al., *A highly specific and discriminatory FISH assay for distinguishing between benign and malignant melanocytic neoplasms*. Am J Surg Pathol, 2012. **36**(6): p. 808-17.
14. Barnhill, R.L., et al., *Atypical Spitz nevi/tumors: lack of consensus for diagnosis, discrimination from melanoma, and prediction of outcome*. Hum Pathol, 1999. **30**(5): p. 513-20.
15. Corona, R., et al., *Interobserver variability on the histopathologic diagnosis of cutaneous melanoma and other pigmented skin lesions*. J Clin Oncol, 1996. **14**(4): p. 1218-23.
16. Farmer, E.R., R. Gonin, and M.P. Hanna, *Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists*. Hum Pathol, 1996. **27**(6): p. 528-31.
17. Lodha, S., et al., *Discordance in the histopathologic diagnosis of difficult melanocytic neoplasms in the clinical setting*. J Cutan Pathol, 2008. **35**(4): p. 349-52.

18. Veenhuizen, K.C., et al., *Quality assessment by expert opinion in melanoma pathology: experience of the pathology panel of the Dutch Melanoma Working Party*. J Pathol, 1997. **182**(3): p. 266-72.
19. Wechsler, J., et al., *Reliability of the histopathologic diagnosis of malignant melanoma in childhood*. Arch Dermatol, 2002. **138**(5): p. 625-8.
20. McArthur, G.A., et al., *Safety and efficacy of vemurafenib in BRAF(V600E) and BRAF(V600K) mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study*. Lancet Oncol, 2014. **15**(3): p. 323-32.
21. Chapman, P.B., et al., *Improved survival with vemurafenib in melanoma with BRAF V600E mutation*. N Engl J Med, 2011. **364**(26): p. 2507-16.
22. Liu, H., et al., *Immunohistochemical detection of the BRAF V600E mutation in melanoma patients with monoclonal antibody VE1*. Pathol Int, 2014. **64**(12): p. 601-6.
23. Ordonez, N.G., *Value of melanocytic-associated immunohistochemical markers in the diagnosis of malignant melanoma: a review and update*. Hum Pathol, 2014. **45**(2): p. 191-205.
24. Painter, J.T., N.P. Clayton, and R.A. Herbert, *Useful immunohistochemical markers of tumor differentiation*. Toxicol Pathol, 2010. **38**(1): p. 131-41.
25. Sheffield, M.V., et al., *Comparison of five antibodies as markers in the diagnosis of melanoma in cytologic preparations*. Am J Clin Pathol, 2002. **118**(6): p. 930-6.
26. Kaudewitz, P., et al., *Tumor cell growth fractions in human malignant melanomas and the correlation to histopathologic tumor grading*. Am J Pathol, 1989. **134**(5): p. 1063-8.
27. Chen, H., et al., *S100 protein family in human cancer*. Am J Cancer Res, 2014. **4**(2): p. 89-115.
28. Gaynor, R., et al., *S100 protein: a marker for human malignant melanomas?* Lancet, 1981. **1**(8225): p. 869-71.
29. Gown, A.M., et al., *Monoclonal antibodies specific for melanocytic tumors distinguish subpopulations of melanocytes*. Am J Pathol, 1986. **123**(2): p. 195-203.
30. Steuhl, K.P., J.M. Rohrbach, and M. Knorr, *[Distribution of melanoma-associated antigens (HMB 45 and S 100) in benign and malignant melanocytic tumors of the conjunctiva]*. Klin Monbl Augenheilkd, 1991. **199**(3): p. 187-91.
31. Levy, C., M. Khaled, and D.E. Fisher, *MITF: master regulator of melanocyte development and melanoma oncogene*. Trends Mol Med, 2006. **12**(9): p. 406-14.
32. Buonaccorsi, J.N., et al., *Diagnostic utility and comparative immunohistochemical analysis of MITF-1 and SOX10 to distinguish melanoma in situ and actinic keratosis: a clinicopathological and immunohistochemical study of 70 cases*. Am J Dermatopathol, 2014. **36**(2): p. 124-30.
33. Fetsch, P.A., et al., *Melanoma-associated antigen recognized by T cells (MART-1): the advent of a preferred immunocytochemical antibody for the diagnosis of metastatic malignant melanoma with fine-needle aspiration*. Cancer, 1999. **87**(1): p. 37-42.
34. Orchard, G.E., *Melan A (MART-1): a new monoclonal antibody for malignant melanoma diagnosis*. Br J Biomed Sci, 1998. **55**(1): p. 8-9.

35. Hu, L., et al., *Fluorescence in situ hybridization (FISH): an increasingly demanded tool for biomarker research and personalized medicine*. *Biomark Res*, 2014. **2**(1): p. 3.
36. North, J.P., et al., *Fluorescence in situ hybridization as an ancillary tool in the diagnosis of ambiguous melanocytic neoplasms: a review of 804 cases*. *Am J Surg Pathol*, 2014. **38**(6): p. 824-31.
37. Gerami, P., et al., *Fluorescence in situ hybridization (FISH) as an ancillary diagnostic tool in the diagnosis of melanoma*. *Am J Surg Pathol*, 2009. **33**(8): p. 1146-56.
38. Pouryazdanparast, P., et al., *Distinguishing epithelioid blue nevus from blue nevus-like cutaneous melanoma metastasis using fluorescence in situ hybridization*. *Am J Surg Pathol*, 2009. **33**(9): p. 1396-400.
39. Requena, C., et al., *Fluorescence in situ hybridization for the differential diagnosis between Spitz naevus and spitzoid melanoma*. *Histopathology*, 2012. **61**(5): p. 899-909.
40. Speicher, M.R. and N.P. Carter, *The new cytogenetics: blurring the boundaries with molecular biology*. *Nat Rev Genet*, 2005. **6**(10): p. 782-92.
41. Ried, T., et al., *Tumor cytogenetics revisited: comparative genomic hybridization and spectral karyotyping*. *J Mol Med (Berl)*, 1997. **75**(11-12): p. 801-14.
42. Hermsen, M.A., et al., *Comparative genomic hybridization: a new tool in cancer pathology*. *Hum Pathol*, 1996. **27**(4): p. 342-9.
43. Bauer, J. and B.C. Bastian, *Distinguishing melanocytic nevi from melanoma by DNA copy number changes: comparative genomic hybridization as a research and diagnostic tool*. *Dermatol Ther*, 2006. **19**(1): p. 40-9.
44. Kirchhoff, M., et al., *Deletions below 10 megabasepairs are detected in comparative genomic hybridization by standard reference intervals*. *Genes Chromosomes Cancer*, 1999. **25**(4): p. 410-3.
45. Chandler, W.M., et al., *Differentiation of malignant melanoma from benign nevus using a novel genomic microarray with low specimen requirements*. *Arch Pathol Lab Med*, 2012. **136**(8): p. 947-55.
46. Minca, E.C., et al., *Genomic microarray analysis on formalin-fixed paraffin-embedded material for uveal melanoma prognostication*. *Cancer Genet*, 2014. **207**(7-8): p. 306-15.
47. Little, S.E., et al., *Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA*. *Genomics*, 2006. **87**(2): p. 298-306.
48. Horsley, S.W., et al., *Array CGH of fusion gene-positive leukemia-derived cell lines reveals cryptic regions of genomic gain and loss*. *Genes Chromosomes Cancer*, 2006. **45**(6): p. 554-64.
49. Ahn, J.W., et al., *Array CGH as a first line diagnostic test in place of karyotyping for postnatal referrals - results from four years' clinical application for over 8,700 patients*. *Mol Cytogenet*, 2013. **6**(1): p. 16.
50. Pinkel, D. and D.G. Albertson, *Array comparative genomic hybridization and its applications in cancer*. *Nat Genet*, 2005. **37 Suppl**: p. S11-7.
51. Shastry, B.S., *SNPs: impact on gene function and phenotype*. *Methods Mol Biol*, 2009. **578**: p. 3-22.

52. Amos, C.I., et al., *Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1*. Nat Genet, 2008. **40**(5): p. 616-22.
53. Tuupanen, S., et al., *The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling*. Nat Genet, 2009. **41**(8): p. 885-90.
54. Kvaskoff, M., et al., *Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma*. Twin Res Hum Genet, 2011. **14**(5): p. 422-32.
55. Holderfield, M., et al., *Targeting RAF kinases for cancer therapy: BRAF-mutated melanoma and beyond*. Nat Rev Cancer, 2014. **14**(7): p. 455-67.
56. Van Raamsdonk, C.D., et al., *Frequent somatic mutations of GNAQ in uveal melanoma and blue naevi*. Nature, 2009. **457**(7229): p. 599-602.
57. Van Raamsdonk, C.D., et al., *Mutations in GNA11 in uveal melanoma*. N Engl J Med, 2010. **363**(23): p. 2191-9.
58. Kawashima, Y., et al., *Efficient extraction of proteins from formalin-fixed paraffin-embedded tissues requires higher concentration of tris(hydroxymethyl)aminomethane*. Clin Proteomics, 2014. **11**(1): p. 4.
59. Ludyga, N., et al., *Nucleic acids from long-term preserved FFPE tissues are suitable for downstream analyses*. Virchows Arch, 2012. **460**(2): p. 131-40.
60. Shibata, D., W.J. Martin, and N. Arnheim, *Analysis of DNA sequences in forty-year-old paraffin-embedded thin-tissue sections: a bridge between molecular biology and classical histology*. Cancer Res, 1988. **48**(16): p. 4564-6.
61. Srinivasan, M., D. Sedmak, and S. Jewell, *Effect of fixatives and tissue processing on the content and integrity of nucleic acids*. Am J Pathol, 2002. **161**(6): p. 1961-71.
62. Bonin, S., et al., *PCR analysis in archival postmortem tissues*. Mol Pathol, 2003. **56**(3): p. 184-6.
63. Tuefferd, M., et al., *Genome-wide copy number alterations detection in fresh frozen and matched FFPE samples using SNP 6.0 arrays*. Genes Chromosomes Cancer, 2008. **47**(11): p. 957-64.
64. Paireder, S., et al., *Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues*. Anal Biochem, 2013. **439**(2): p. 152-60.
65. Wang, J.H., et al., *DNA extraction from fresh-frozen and formalin-fixed, paraffin-embedded human brain tissue*. Neurosci Bull, 2013. **29**(5): p. 649-54.
66. Rabelo-Goncalves, E., et al., *Evaluation of five DNA extraction methods for detection of H. pylori in formalin-fixed paraffin-embedded (FFPE) liver tissue from patients with hepatocellular carcinoma*. Pathol Res Pract, 2014. **210**(3): p. 142-6.
67. Heydt, C., et al., *Comparison of pre-analytical FFPE sample preparation methods and their impact on massively parallel sequencing in routine diagnostics*. PLoS One, 2014. **9**(8): p. e104566.
68. van Eijk, R., et al., *Assessment of a fully automated high-throughput DNA extraction method from formalin-fixed, paraffin-embedded tissue for KRAS, and BRAF somatic mutation analysis*. Exp Mol Pathol, 2013. **94**(1): p. 121-5.

69. Ling Lin, et al. *Improved FFPE DNA extraction for next generation sequencing using adaptive focused acoustics technology*. in *105th Annual Meeting of the American Association for Cancer Research*. 2014. San Diego, CA: Cancer Research.
70. Nelson, D.S., et al., *Somatic activating ARAF mutations in Langerhans cell histiocytosis*. *Blood*, 2014. **123**(20): p. 3152-5.
71. Yap, K.L., et al., *Whole-Exome Sequencing of Muscle-Invasive Bladder Cancer Identifies Recurrent Mutations of UNC5C and Prognostic Importance of DNA Repair Gene Mutations on Survival*. *Clin Cancer Res*, 2014. **20**(24): p. 6605-17.
72. I, A., *Genotyping Console 4.0 User Manual*. 2009
73. Anderson, C.A., et al., *Data quality control in genetic case-control association studies*. *Nat Protoc*, 2010. **5**(9): p. 1564-73.
74. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses*. *Am J Hum Genet*, 2007. **81**(3): p. 559-75.
75. Pathologists, C.o.A., *CAP Accreditation Program Anatomic Pathology Checklist*. 2014, Northfield, IL 60093-2750: College of American Pathologists.
76. Cameron, D., et al., *A graph-based recovery and decomposition of Swanson's hypothesis using semantic predications*. *J Biomed Inform*, 2013. **46**(2): p. 238-51.
77. Ben-Ezra, J., et al., *Effect of fixation on the amplification of nucleic acids from paraffin-embedded material by the polymerase chain reaction*. *J Histochem Cytochem*, 1991. **39**(3): p. 351-4.
78. Isola, J., et al., *Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples*. *Am J Pathol*, 1994. **145**(6): p. 1301-8.
79. Senguven, B., et al., *Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues*. *Int J Med Sci*, 2014. **11**(5): p. 494-9.
80. Specht, K., et al., *Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue*. *Am J Pathol*, 2001. **158**(2): p. 419-29.
81. Karumuri, A.K., et al., *Silver nanoparticles attached to porous carbon substrates: robust materials for chemical-free water disinfection*. *Materials Letters*, 2013. **109**: p. 83-87.
82. Gilbert, M.T., et al., *The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when?* *PLoS One*, 2007. **2**(6): p. e537.
83. von Ahlfen, S., et al., *Determinants of RNA quality from FFPE samples*. *PLoS One*, 2007. **2**(12): p. e1261.
84. Evers, D.L., et al., *The effect of formaldehyde fixation on RNA: optimization of formaldehyde adduct removal*. *J Mol Diagn*, 2011. **13**(3): p. 282-8.
85. Karumuri, A.K., et al., *Fabrication and characterization of antibacterial nanoparticles supported on hierarchical hybrid substrates*. *Journal of Nanoparticle Research*, 2014. **16**(4): p. 1-14.
86. Rhee, H.S. and B.F. Pugh, *ChIP-exo: A Method to Identify Genomic Location of DNA-binding proteins at Near Single Nucleotide Accuracy*. *Current protocols in molecular*

- biology / edited by Frederick M. Ausubel ... [et al.], 2012. **0 21**: p. 10.1002/0471142727.mb2124s100.
87. Goelz, S.E., S.R. Hamilton, and B. Vogelstein, *Purification of DNA from formaldehyde fixed and paraffin embedded human tissue*. Biochem Biophys Res Commun, 1985. **130**(1): p. 118-26.
 88. Stanta, G. and C. Schneider, *RNA extracted from paraffin-embedded human tissues is amenable to analysis by PCR amplification*. Biotechniques, 1991. **11**(3): p. 304, 306, 308.
 89. Simbolo, M., et al., *DNA qualification workflow for next generation sequencing of histopathological samples*. PLoS One, 2013. **8**(6): p. e62692.
 90. Vinh Nguyen, O.B., Jagannathan Srinivasan, Todd Minning, Thomas Rindesch, Bastien Rance, Ramakanth Kavuluru, Hima Yalamanchili , Krishnaprasad Thirunarayan, Satya Sahoo, and Amit Sheth, *iExplore: Interactive Browsing and Exploring Biomedical Knowledge. Semantic Web Challenge*. ISWC, 2012.
 91. Siwoski, A., et al., *An efficient method for the assessment of DNA quality of archival microdissected specimens*. Mod Pathol, 2002. **15**(8): p. 889-92.
 92. Savage, S.J. and G. Hostetter, *Genomic analysis by oligonucleotide array Comparative Genomic Hybridization utilizing formalin-fixed, paraffin-embedded tissues*. Methods Mol Biol, 2011. **700**: p. 185-98.
 93. Craddock, K.J., W.L. Lam, and M.S. Tsao, *Applications of array-CGH for lung cancer*. Methods Mol Biol, 2013. **973**: p. 297-324.
 94. Dimova, I., et al., *Genomic markers for ovarian cancer at chromosomes 1, 8 and 17 revealed by array CGH analysis*. Tumori, 2009. **95**(3): p. 357-66.
 95. Ishkanian, A.S., et al., *High-resolution array CGH identifies novel regions of genomic alteration in intermediate-risk prostate cancer*. Prostate, 2009. **69**(10): p. 1091-100.
 96. Trolet, J., et al., *Genomic profiling and identification of high-risk uveal melanoma by array CGH analysis of primary tumors and liver metastases*. Invest Ophthalmol Vis Sci, 2009. **50**(6): p. 2572-80.
 97. Beroukhim, R., et al., *Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma*. Proc Natl Acad Sci U S A, 2007. **104**(50): p. 20007-12.
 98. Nair, B.C., et al., *Roscovitine confers tumor suppressive effect on therapy-resistant breast tumor cells*. Breast Cancer Res, 2011. **13**(3): p. R80.
 99. Maryam Panahiazar, A.R., Vahid Taslimi, Hima Yalamanchili, Arlin Stoltzfus, Jim Leebens-Mack, and Amit Sheth, *PhylOnt : A Domain-Specific Ontology for Phylogeny Analysis IEEE International Conference on Bioinformatics and Biomedicine*. BIBM, 2012. **52: 1, 106-116**
 100. Turner, S., et al., *Quality control procedures for genome-wide association studies*. Curr Protoc Hum Genet, 2011. **Chapter 1**: p. Unit1 19.
 101. Stefanaki, I., et al., *Replication and predictive value of SNPs associated with melanoma and pigmentation traits in a Southern European case-control study*. PLoS One, 2013. **8**(2): p. e55712.
 102. Iles, M.M., et al., *The effect on melanoma risk of genes previously associated with telomere length*. J Natl Cancer Inst, 2014. **106**(10).

103. Meyer, P., C. Sergi, and C. Garbe, *Polymorphisms of the BRAF gene predispose males to malignant melanoma*. J Carcinog, 2003. **2**(1): p. 7.
104. Gu, F., et al., *Interleukin and interleukin receptor gene polymorphisms and susceptibility to melanoma*. Melanoma Res, 2008. **18**(5): p. 330-5.
105. Maccioni, L., et al., *Variants at chromosome 20 (ASIP locus) and melanoma risk*. Int J Cancer, 2013. **132**(1): p. 42-54.
106. Rafnar, T., et al., *Sequence variants at the TERT-CLPTM1L locus associate with many cancer types*. Nat Genet, 2009. **41**(2): p. 221-7.
107. Vinagre, J., et al., *Frequency of TERT promoter mutations in human cancers*. Nat Commun, 2013. **4**: p. 2185.
108. Williams, N.G. and T.M. Roberts, *Signal transduction pathways involving the Raf proto-oncogene*. Cancer Metastasis Rev, 1994. **13**(1): p. 105-16.
109. Vallabhaneni, S., et al., *Significance of ER-Src axis in hormonal therapy resistance*. Breast Cancer Res Treat, 2011. **130**(2): p. 377-85.
110. Uribe, P., Wistuba, II, and S. Gonzalez, *BRAF mutation: a frequent event in benign, atypical, and malignant melanocytic lesions of the skin*. Am J Dermatopathol, 2003. **25**(5): p. 365-70.
111. Suzuki, I., et al., *Agouti signaling protein inhibits melanogenesis and the response of human melanocytes to alpha-melanotropin*. J Invest Dermatol, 1997. **108**(6): p. 838-42.
112. Wan, M., et al., *OB fold-containing protein 1 (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation*. J Biol Chem, 2009. **284**(39): p. 26725-31.
113. Lake, S.L., et al., *Single nucleotide polymorphism array analysis of uveal melanomas reveals that amplification of CNKSR3 is correlated with improved patient survival*. Am J Pathol, 2013. **182**(3): p. 678-87.
114. Nikolaev, S.I., et al., *Exome sequencing identifies recurrent somatic MAP2K1 and MAP2K2 mutations in melanoma*. Nat Genet, 2012. **44**(2): p. 133-9.
115. Shimizu, K., et al., *Characterization of phosphodiesterase 1 in human malignant melanoma cell lines*. Anticancer Res, 2009. **29**(4): p. 1119-22.
116. Perrot, C.Y., D. Javelaud, and A. Mauviel, *Insights into the Transforming Growth Factor-beta Signaling Pathway in Cutaneous Melanoma*. Ann Dermatol, 2013. **25**(2): p. 135-44.
117. Ellmann, L., et al., *BRN2 is a transcriptional repressor of CDH13 (T-cadherin) in melanoma cells*. Lab Invest, 2012. **92**(12): p. 1788-800.
118. Prickett, T.D., et al., *Somatic mutation of GRIN2A in malignant melanoma results in loss of tumor suppressor activity via aberrant NMDAR complex formation*. J Invest Dermatol, 2014. **134**(9): p. 2390-8.
119. Pauli, B.U., et al., *Molecular characteristics and functional diversity of CLCA family members*. Clin Exp Pharmacol Physiol, 2000. **27**(11): p. 901-5.
120. Arai, E., et al., *Multilayer-omics analysis of renal cell carcinoma, including the whole exome, methylome and transcriptome*. Int J Cancer, 2014. **135**(6): p. 1330-42.
121. Wang, W., et al., *ERK3 promotes endothelial cell functions by upregulating SRC-3/SP1-mediated VEGFR2 expression*. J Cell Physiol, 2014. **229**(10): p. 1529-37.

122. Savai, R., et al., *Targeting cancer with phosphodiesterase inhibitors*. *Expert Opin Investig Drugs*, 2010. **19**(1): p. 117-31.
123. Han, S.U., et al., *Loss of the Smad3 expression increases susceptibility to tumorigenicity in human gastric cancer*. *Oncogene*, 2004. **23**(7): p. 1333-41.
124. Sataranatarajan, K., et al., *Molecular events in matrix protein metabolism in the aging kidney*. *Aging Cell*, 2012. **11**(6): p. 1065-73.
125. Hazan, R.B., et al., *Cadherin switch in tumor progression*. *Ann N Y Acad Sci*, 2004. **1014**: p. 155-63.
126. Konieczkowski, D.J., et al., *A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors*. *Cancer Discov*, 2014. **4**(7): p. 816-27.
127. Kalkman, H.O., *Potential opposite roles of the extracellular signal-regulated kinase (ERK) pathway in autism spectrum and bipolar disorders*. *Neurosci Biobehav Rev*, 2012. **36**(10): p. 2206-13.
128. Ding, L., et al., *Clonal architectures and driver mutations in metastatic melanomas*. *PLoS One*, 2014. **9**(11): p. e111153.
129. Gruber, A.D. and B.U. Pauli, *Tumorigenicity of human breast cancer is associated with loss of the Ca²⁺-activated chloride channel CLCA2*. *Cancer Res*, 1999. **59**(21): p. 5488-91.
130. Riker, A.I., et al., *The gene expression profiles of primary and metastatic melanoma yields a transition point of tumor progression and metastasis*. *BMC Med Genomics*, 2008. **1**: p. 13.
131. Khalili, J.S., et al., *Combination small molecule MEK and PI3K inhibition enhances uveal melanoma cell death in a mutant GNAQ- and GNA11-dependent manner*. *Clin Cancer Res*, 2012. **18**(16): p. 4345-55.
132. Potluri, K., et al., *Genomic DNA extraction methods using formalin-fixed paraffin-embedded tissue*. *Anal Biochem*, 2015. **486**: p. 17-23.
133. Johnson, N.A., et al., *Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells*. *Lab Invest*, 2006. **86**(9): p. 968-78.
134. Azimzadeh, O., et al., *Formalin-fixed paraffin-embedded (FFPE) proteome analysis using gel-free and gel-based proteomics*. *J Proteome Res*, 2010. **9**(9): p. 4710-20.
135. Giusti, L. and A. Lucacchini, *Proteomic studies of formalin-fixed paraffin-embedded tissues*. *Expert Rev Proteomics*, 2013. **10**(2): p. 165-77.
136. Edwin Rudd, J.H., Austin Purdy, Hamid Khoja, J.D. Herlihy, Guillaume Durin, *DNA Extraction from FFPE samples with Covaris Focused-ultrasonicators*. Covaris, Inc., Woburn, Massachusetts, USA, 2013.
137. Salawu, A., et al., *High quality genomic copy number data from archival formalin-fixed paraffin-embedded leiomyosarcoma: optimisation of universal linkage system labelling*. *PLoS One*, 2012. **7**(11): p. e50415.
138. Orlow, I., et al., *Genetic factors associated with naevus count and dermoscopic patterns: preliminary results from the Study of Nevi in Children (SONIC)*. *Br J Dermatol*, 2015. **172**(4): p. 1081-9.