

바이오 인포매틱스

졸업 논문

개인 유전체 분석을 통한
metabolism 형질 분석

바이오 인포매틱스 연계전공

1748039 백나림

2021.5.27

목차

서론	3
연구방법	4
Exome sequencing	4
Raw data analysis	4
NCAD metabolism Catalog	5
연구결과 및 토의	6
Nicotine	6
Caffeine	9
Alcohol	10
Drugs	12
결론	16
참고문헌	17
ACKNOWLEDGEMENT	19

서론

인간의 유전자는 약 22,000 개 이상이 존재하며 엑손(exon) 부위는 단백질을 직접 코딩하는 부위로 전체 유전체의 1-2%정도를 차지하고 있다. 대부분의 질환 연관 돌연변이는 엑손에 존재하기 때문에 Exome sequencing 을 통해 효율적으로 돌연변이를 검출할 수 있다. 개인 유전체 분석을 통해 보편적인 reference genome(=ref gene, 참조 염기서열)과 비교해보면 유전자가 어떤 variants 를 가지는지 알 수 있고, 이것으로 개인이 어떤 특성을 갖고 있는지 예측할 수 있다. 또한 질병이 없는 사람들과 질병을 가진 사람의 전체 유전체를 대조해 보면, 앞으로 발생 가능성이 큰 pathogenicity 도 알 수 있어 여러 질병을 예측하고 진단하는 데 도움을 줄 수 있다.

본 연구에서는 건강한 성인 6 명의 상피세포를 채취하여 exome sequencing 을 진행한 raw data 를 germline variant call pipeline 에 따라 가공하였고 ref gene 으로는 GRCh38(hg38) 버전을 선택하였다. 이를 위해 sickle 로 염기서열 절편인 read 를 trimming, ref gene 에 indexing, bowtie2 로 alignment, samtools 로 파일 형식을 변환하였다. 그리고 germline variant caller 인 gatk-haplotype caller 와 strelka 를 이용하여 필터를 설정하고 자료를 추출하여 VEP 와 annovar 로 annotation 해주었다. 이후에 6 명 간 공통적으로 가지는 SNP(Single Nucleotide Polymorphism)를 확인하고 어떤 특성이 있는지 연관 분석하였다.

6 명이 가진 SNP 의 특성을 확인하기 위해 GWAS 를 이용하면서 SNP 와 형질과의 인과관계가 구체적으로 명시되지 않아 이를 보완할 필요성을 느꼈다. 다양한 형질 중에서 일상생활에서 특히 섭취 빈도가 높은 니코틴, 카페인, 술, 의약품의 metabolism 과 관련된 정보들을 GWAS 에서 선별하였고, 해당 SNP 에 따라 어떤 형질을 가질 수 있는지를 구체적으로 명시한 'NCAD(Nicotine, Caffeine, Alcohol, Drugs)-metabolism catalog'를 만들어 보았다. 그리고 이를 앞서 만든 6 명의 개인 유전체 자료와 비교하여 유전적 변이에 기반한 예측과 실제 개인의 특성이 일치하는지 확인하는 과정을 거쳤다.

연구 방법

1. 엑솜 시퀀싱(Whole exome Sequencing)

인간의 유전자는 약 22,000 개 이상이 존재하며 엑손(exon) 부위는 단백질을 직접 코딩하는 부위로 전체 유전체의 1-2%정도를 차지하고 있다. 대부분의 질환 연관 돌연변이는 엑손에 존재하기 때문에 엑솜 시퀀싱을 통해 효율적으로 돌연변이를 검출할 수 있다. 엑솜 시퀀싱은 중간 정도의 시퀀싱 깊이를 얻을 수 있고 유전체 전체를 분석하는 전장 유전체 시퀀싱(Whole genome sequencing)에 비해서 분석에 소요되는 시간이 줄어들고 비용이 저렴하여 효율적이다.¹ 본 연구에서는 건강한 성인 6 명의 상피세포를 채취하여 exome sequencing 을 진행하였다.

2. Raw-data Analysis

시퀀싱 결과 생성된 FASTQ 파일의 짧은 리드들이 어떤 염색체의 어느 위치에 있는지 찾아주기 위해 인간의 표준 유전체로 GRCh38(hg38) 버전을 선택하였다. 이를 위해 우선 sickle 로 염기서열 절편인 read 를 trimming 한 후, ref gene 에 indexing, bowtie2 로 alignment 하여 samtools 로 파일 형식을 SAM/BAM 파일로 변환하였다. 매핑 이후 각 시퀀싱 리드를 분석하여 특정 위치에서 표준 유전체 서열과 다른 변이(variation)를 찾아내기 위해 germline variant caller 인 gatk-haplytype caller 와 strelka 를 이용하여 필터를 설정하고 자료를 추출한 후 VEP 과 Annovar 로 genome 의 변이에 대한 database 정보를 annotation 하는 작업을 진행하였다.

¹ 이승태, Next Generation Sequencing 기반 유전자 검사의 이해, 식품의약품안전처, p.19

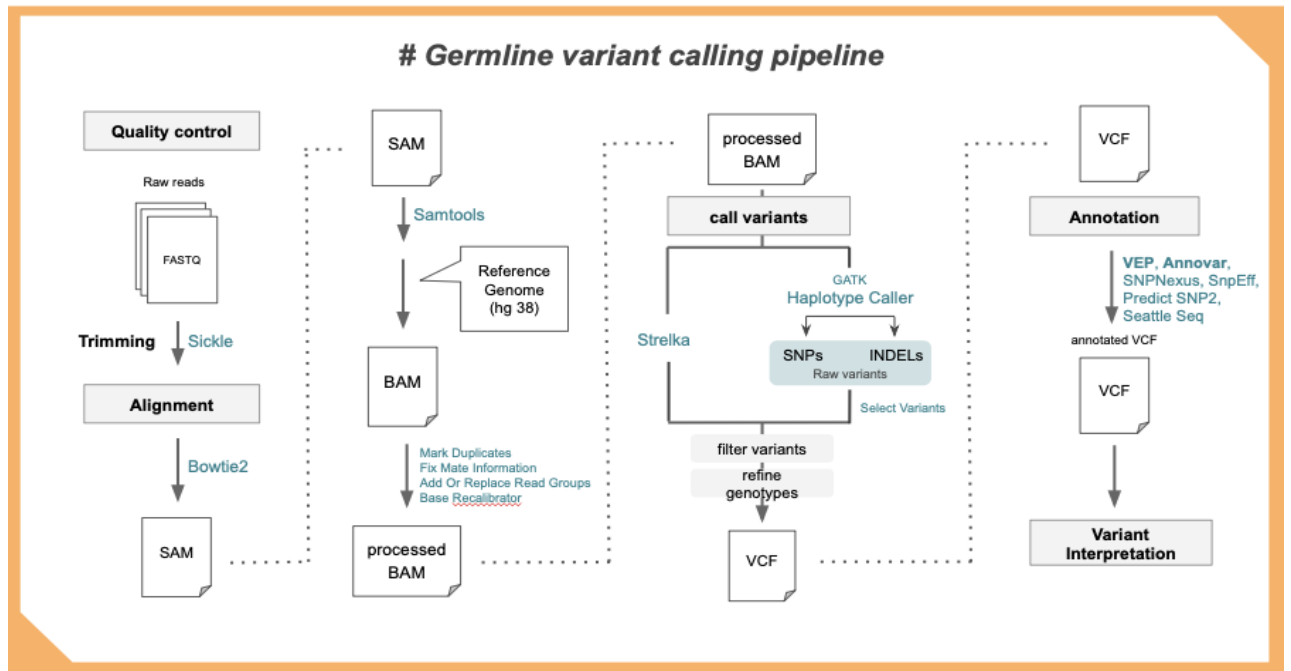


Figure1. germline variant calling pipeline

3. NCAD Metabolism Catalog

유전체 변이 정보를 찾기 위해 GWAS 를 이용하면서 SNP 와 형질과의 인과관계가 구체적으로 명시되지 않아 이를 보완할 필요성을 느꼈다. 이에 따라 일상생활에서 특히 섭취 빈도가 높은 니코틴(Nicotine), 카페인(Caffeine), 알코올(Alcohol), 의약품(Drugs)의 metabolism 카테고리를 만들어서 해당하는 형질에 관여하는 유전자 정보를 GWAS 에서 선별하여 구체적으로 명시한 NCAD-metabolism-catalog 을 제작하였다. 그리고 이를 앞서 만든 6 명의 개인 SNP 자료와 비교하여 예측 모델과 실제 개인의 특성이 일치하는지 확인하였다.

연구 결과 및 토의

1. 니코틴(Nicotine)

흡연은 유전적 요인에 따라 질병 발생의 위험도가 달라진다고 알려져 있다. 특히 인종에 따라 니코틴의 대사 비율이 달라지기 때문에 니코틴의 생체 축적이 다르게 일어난다는 연구도 보고되고 있다.² 니코틴의 대사 과정은 대체로 Figure2 와 같다. 니코틴은 cytochrome P-450(cyp) 효소작용에 의해 C-oxidation, N-oxidation 반응이 일어나 cotinine 과 nicotine-1-N-oxide 로 되며, 일부는 N-glucuronide 로 대사된다. cotinine 은 다시 대사되어 trans-3-hydroxycotinine, cotinine-glucuronide, trans-3-hydroxycotinine glucuronide 등으로 진행된다.³ 이때 cotinine 이 nicotine 의 대사에서 약 80%를 차지하는 주된 대사산물로, 이 과정에서 CYP2A6 이라는 효소가 관여한다. 니코틴의 대사 정도를 판단하기 위한 지표로 NMR 이라 불리는 니코틴 대사 비율은 hydroxycotinine 대 cotinine 비율을 의미하며 NMR 값이 클수록 nicotine 대사 속도가 빠르다고 판단된다. 체내에 동일한 nicotine 수준을 유지하기 위하여 nicotine 대사가 빠르게 일어나는 사람들은 상대적으로 더 많은 담배를 피는 경향이 있고 결과적으로 담배 연기에 노출되는 빈도가 잦아 폐암의 위험성이 높다고 보고되고 있다.

² Yerger, V.B., Malone, R.E. Melanin and nicotine: A review of the literature. Nicotine & Tobacco Research. 2006, 8(4), 487-498. doi:10.1080/14622200600790039

³ 권준택. "니코틴 대사와 CYP2A6 유전자 빈도에 관한 민족적 다형성." Journal of Soonchunhyang Medical Science 6.2 (2000): 285-292.

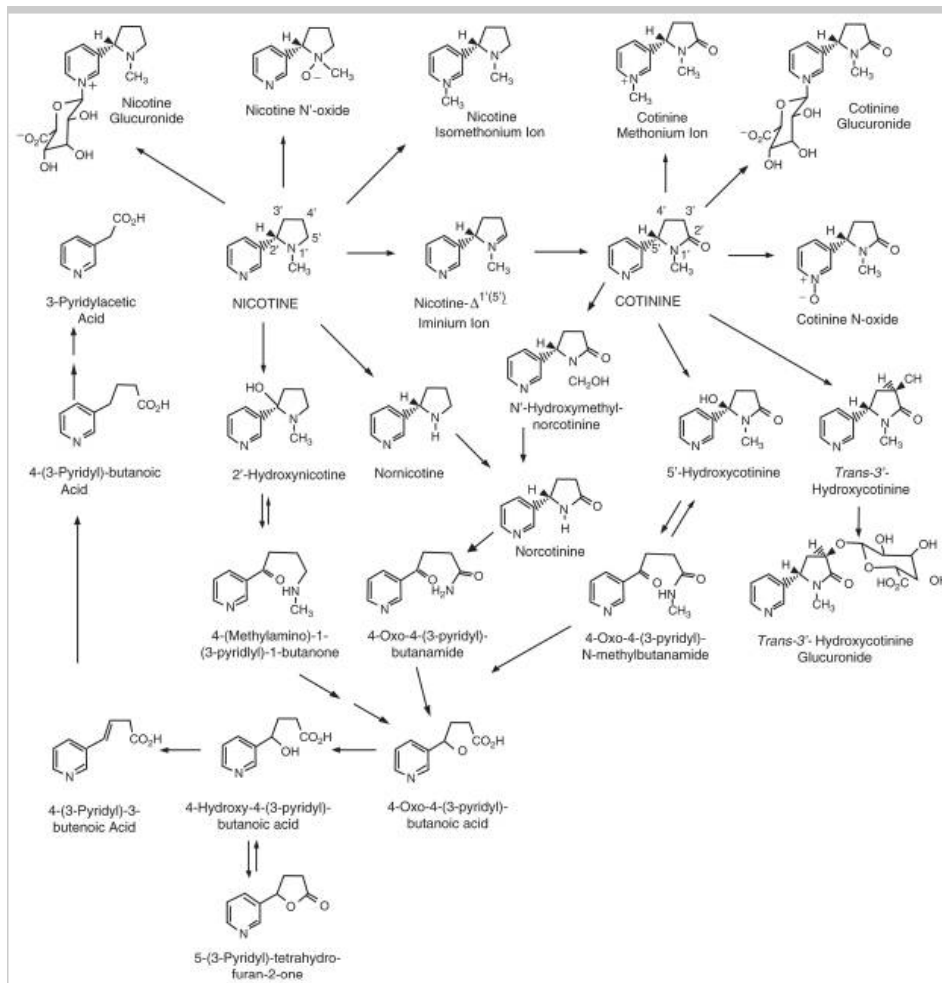


Figure2. Metabolism of Nicotine⁴

CYP2A6 에 의한 니코틴 대사는 흡연 용량과 폐암 위험성을 결정하는 것으로 의심되는 요인이다. GWAS 에서 조사한 결과 니코틴 대사와 관련된 18 개의 변이를 확인할 수 있었으며 대부분 19q13.2 위치에 존재하였으며 CYP2A6, CYP2A7, EGLN2 유전자 변이인 경우가 많았다. 조사한 snp 를 가진 사람들은 니코틴 대사가 상대적으로 빠르게 일어나서 체내 니코틴 비율을 일정 정도로 유지시키기 위하여 니코틴 소비량, 의존성, 관련 질병 위험성이 비교적 높게 나타날 것으로 예측된다. 조사한 6 명은 모두 비흡연자였으며, 개인 유전체 자료와 제작한 데이터베이스를 비교 분석한 결과 니코틴 대사와 관련된 변이는 6 명 모두 나타나지 않았다.

⁴ Hukkanen J, Jacob P, 3rd, Benowitz NL. Metabolism and disposition kinetics of nicotine. *Pharmacol Rev.* 2005c;57(1):79-115

PUBMED ID	DISEASE/TRAIT	REGION	CHR_ID	CHR_POS	REPORTED GENE(S)	MAPPED_GENE	STRONGEST SNP-RISK ALLELE	CONTEXT	RISK ALLELE FREQUENCY	P-VALUE	MAPPED TRAIT
29460428	Response to alcohol consumption (flushing response 흥조)	12q24.13	12	112469070	ALDH2, PTPN11, NUP-PTPN11	AC010967.1 - AC069157.2	rs143894582-C	intron_variant	0.9195	2E-14	response to alcohol
29460428	Response to alcohol consumption (flushing response)	2p16.2	2	5326605	intergenic	NAALADL2	rs200948948-G	intergenic_variant	0.464	4E-07	response to alcohol
29460428	Response to alcohol consumption (flushing response)	3q26.31	3	175318979	NAALADL2	NAALADL2	rs397813807-A	intron_variant	0.8135	4E-06	response to alcohol
29460428	Response to alcohol consumption (flushing response)	6q22.33	6	12957556	intergenic	AL36124.1	rs10457526-G	intron_variant	0.7298	5E-06	response to alcohol
29460428	Response to alcohol consumption (flushing response)	7q31.2	7	117175543	ST7	ST7	rs14335618-C	intron_variant	0.114	3E-06	response to alcohol
29460428	Response to alcohol consumption (flushing response)	8p11.21	8	40280820	intergenic	AC02733.2 - SJRLNT	rs2925009-C	intergenic_variant	0.6463	3E-06	response to alcohol
29460428	Response to alcohol consumption (flushing response)	11p15.5	11	1534014	intergenic, DUSP8	MOB2 - DUSP8	rs578492-A	intergenic_variant	0.5794	4E-06	response to alcohol
29460428	Response to alcohol consumption (flushing response)	12q24.12	12	111803962	ALDH2	ALDH2	rs671-?	missense_variant			response to alcohol
24277619	Response to alcohol consumption (flushing response)	14q24.2	14	71520348	intergenic	SIPA1L1	rs11158907-?	intron_variant			response to alcohol
24277619	Response to alcohol consumption (flushing response)	8q24.22	8	134866016	intergenic	AC103764.1	rs4909801-?	intron_variant			response to alcohol
30276832	Initial alcohol sensitivity	1q44	1	244399651	Clorf100	AL358177.1 - Clorf100	rs145005509-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	1q44	1	247912096	OR2T8	OR2W3 - OR2T8	rs10788734-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	2p16.1	2	59274730	ENSG00000233891	LINC01793, AC007179.2, rs172806266-?	AC012512.1 - AP153	regulatory_region_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	2q36.1	2	22373478	AP153	AC012512.1 - AP153	rs112834343-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	3p25.3	3	10872053	SLC6A11	SLC6A11	rs17033567-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	3p24.3	3	21982028	ZNF850-AS2	ZNF850	rs233622-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	3q22.1	3	133499064	ENSG00000214301	AC022296.2 - AC022296.4	rs112368179-?	regulatory_region_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	4p16.3	4	542637	PIGG	PIGG - TMEM271	rs75536499-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	4q21.23	4	85432557	ARHGAP24	AC097488.1 - ARHGAP24	rs115496994-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	4q35.1	4	183250034	WWC2	WWC2	rs10020261-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	5q15	5	96315649	CTD-237A12.1	AC104123.1	rs486981-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	6p12.3	6	51060459	FTHP5	FTHP5 - AL158050.2	rs75886551-?	regulatory_region_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	6p12.2	6	52243786	ILLTF	ILLTF	rs11465543-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	6q15	6	8756714	RARS2	RARS2	rs7656242-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	6q21	6	113333595	ENSG00000223811	SOC5P5 - AL589684.1	rs62421504-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	6q27	6	167276064	UNC93A	UNC93A	rs206972-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	7p11.2	7	55051808	EGFR	EGFR	rs73133463-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	8p23.3	8	477140	ENSG00000272005	FBXQ25	rs2100160-?	3_prime_UTR_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	8q24.22	8	133893495	RP11-157E21.1	AC110741.1	rs16905012-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	8q24.23	8	137533321	ENSG00000254076	AC105213.1	rs11777857-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	9q34.3	9	137103590	MAN1B1	MAN1B1	rs28373932-?	non_coding_transcript_ew	NR		response to alcohol
30276832	Initial alcohol sensitivity			RSU1	RSU1		rs76238752-?		NR		response to alcohol
30276832	Initial alcohol sensitivity	10q21.1	10	54833105	PCDH15	PCDH15	rs10825405-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	10q21.3	10	65534026	ENSG00000228065	AL592466.1 - LINC01515	rs75752490-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	10q23.1	10	83922871	ENSG00000232258	AL390786.1 - RNU1-65P	rs61866256-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	10q24.31	10	100108990	TPM4P1	CYP2C23P	rs707625-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	10q25.1	10	108019723	RP11-215N21.1	LINC01435	rs184338590-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	11q13.4	11	71358209	AP002387.1	SHANK2 - AP002387.1	rs75794081-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	13q31.1	13	85943505	SLITRK6	AL162373.1 - MOB1JAP1	rs9547398-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	14q12	14	26210194	CYB5AP5	CYB5AP5 - NOVA1	rs1016246-?	intergenic_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	15q21.1	15	45200754	SUF	SUF	rs116879015-?	5_prime_UTR_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	16q21	16	57900176	CNGB1	CNGB1	rs146087183-?	intron_variant	NR		response to alcohol
30276832	Initial alcohol sensitivity	17p13.2	17	4775560	TN45F5	TN45F5	rs7214066-?	intron_variant	NR		response to alcohol

Figure3. NACD-metabolism-catalog 일부

2. 카페인(Caffeine)

카페인은 전세계적으로 소비량이 가장 높은 향정신성 물질로 대사 과정에 있어서 개인간에 다양한 변화를 가져온다. 카페인을 섭취 후 간에서 Paraxanthine(1,7 dimethylxanthine [17X]), theophylline(1,3dimethylxanthine [13X]), theobromine(3,7 dimethylxanthine [37X]) 대사 산물로 주로 분해되며 이 중 CYP1A2 에 의해 분해가 되는 paraxanthine(17X)은 카페인 대사산물에서 80%정도를 차지한다. Paraxanthine 대 caffeine 의 비율(paraxanthine to caffeine ratio)은 혈장 카페인(plasma caffeine)과 함께 카페인 대사 속도를 나타내는 대표적인 지표 중 하나로 사용되는데 해당 비율이 작을수록 카페인 대사가 느리게 일어난다고 판단되며 카페인 대사가 느리게 일어나는 경우는 더 낮은 커피를 소비와 관련이 있다는 사실이 보고되고 있다.

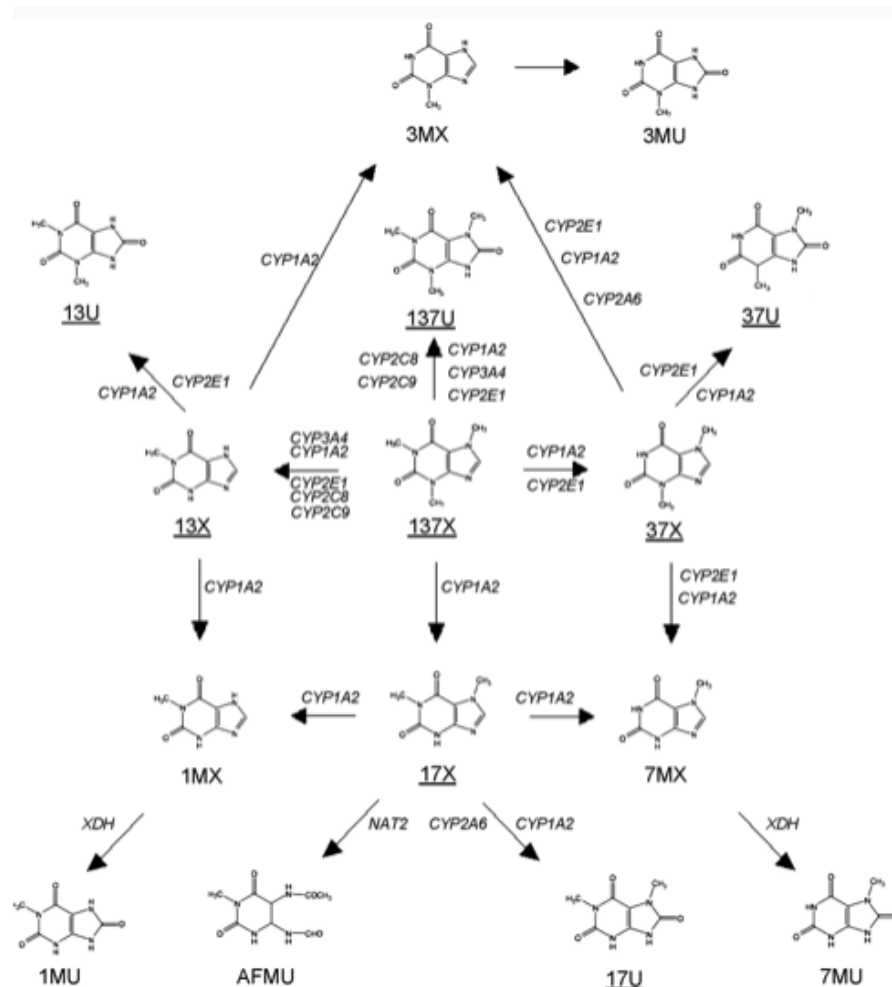


Figure4. Metabolism of Caffeine⁵

⁵ Cornelis, Marilyn C., et al. "Genome-wide association study of caffeine metabolites provides new insights to caffeine metabolism and dietary caffeine-consumption behavior." *Human molecular genetics* 25.24 (2016): 5472-5482.

GWAS 에서 조사한 결과 카페인 대사와 관련된 48 개의 변이를 확인할 수 있었으며 CYP1A2, CYP1A1 유전자 변이인 경우가 많았다. 관련 snp 를 가진 사람들은 카페인 대사가 상대적으로 느리게 일어나서 커피 소비량이 비교적 낮게 나타날 것으로 예측된다. 조사한 6 명 중 1 명은 커피 소비량이 상대적으로 낮았으며, 개인 유전체 자료를 catalog 와 비교 분석한 결과 실제로 커피 소비량이 낮은 1 명을 포함한 총 2 명이 rs1260326-T 변이를 가지고 있었다. 변이가 나타나지 않은 나머지 4 명은 모두 카페인 소비가 높았다.

3. 알코올(Alcohol)

알코올은 우리 일상생활과 쉽게 접할 수 있는 기호 식품으로 1980 년대 이후로 알코올 소비량이 지속적으로 증가하는 추세를 보인다. 알코올의 중간 대사 산물은 여러 생리작용의 변화를 가져와서 다양한 대사성 질환과 간 질환을 가져오는데 이때 알코올에 의한 질환들은 개인의 환경적, 유전적 요인에 따라서 달라진다. 체내에 흡수된 알코올의 일부는 위장에서 대사되고 대부분은 간으로 들어온다. 간으로 들어온 알코올은 세포질에 존재하는 alcohol dehydrogenase(ADH)에 의해 acetaldehyde 로 산화되고, 다시 aldehyde dehydrogenase(ALDH)에 의해 acetate 로 산화된다. 대부분의 acetate 는 acetyl-CoA 로 전환된 후 TCA 회로를 거쳐 에너지를 발생하게 되고, 나머지는 지방산 합성 경로를 거쳐 중성지방으로 합성되어 축적된다⁶. 알코올의 독성은 대사되는 과정 중에서 생성된 acetaldehyde 에 의한 것으로 알려져 있다. 알코올 분해효소라 불리는 aldehyde dehydrogenase(ALDH) 수치가 낮으면 Acetate 로 산화되지 못한 Acetaldehyde 는 체내에 쌓이게 되고 얼굴이 붉어지는 flushing 현상이 나타난다.

⁶ 최지은, "페피노 추출액이 흰쥐의 알코올 대사에 미치는 영향," 국내석사학위논문 단국대학교, 2008. 경기도

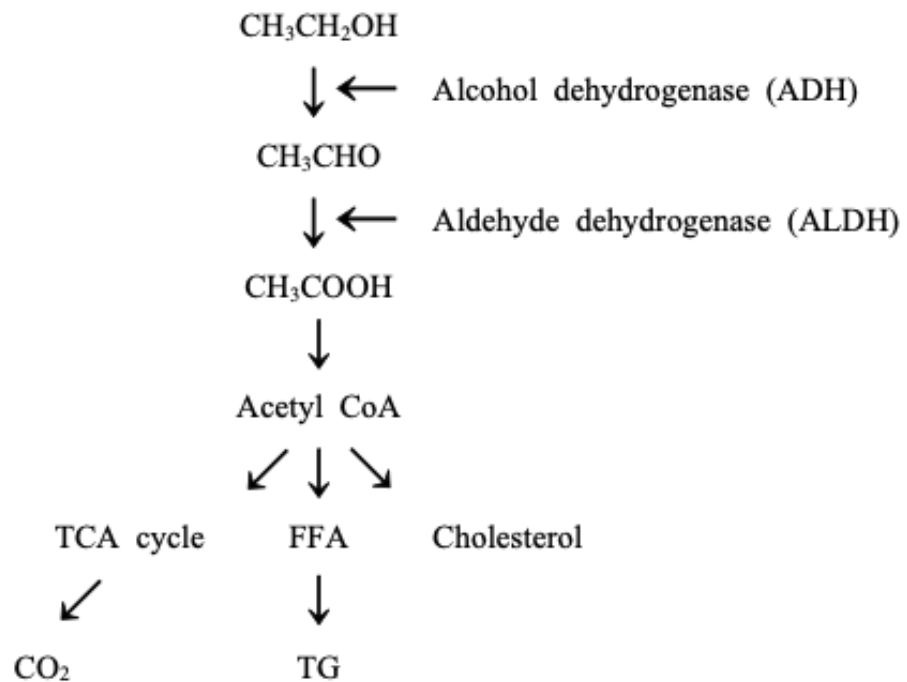


Figure5. Metabolism of alcohol in liver⁷

GWAS 에서 조사한 결과 알코올 대사와 관련된 47 개의 변이를 확인할 수 있었으며 ALDH2 근처 유전자 변이인 경우가 많았다. 관련 snp 를 가진 사람들은 알코올 대사가 상대적으로 느리게 일어나서 체내에 독성을 가진 acetaldehyde 가 비교적 쌓이기 쉽다. 6 명의 개인 유전체 자료와 제작한 데이터베이스를 비교 분석한 결과 알코올 대사와 관련된 변이 rs671-T 를 가진 3 명이 존재하였다. 실제로 변이를 가지고있는 3 명은 음주 후 얼굴이 붉어지는 flushing 이 나타나며, 알코올 소비량도 상대적으로 낮게 나타났다. 알코올 대사와 관련된 변이가 나타나지 않은 나머지 3 명은 음주 후 flushing 현상이 나타나지 않았다.

⁷ Linder MC. (1985). Nutritional biochemistry and metabolism. Elsevier, p. 48

4. Drugs

약물은 환자의 증상을 완화하고 병의 재발 빈도를 줄이는데 도움을 주지만, 치료 효과나 부작용은 환자의 유전적, 환경적 요인에 따라 다양하게 나타난다. GWAS 에서 tamsulosin, opiate, quetiapine, thiopurine, acetaminophen, warfarin, methylphenidate, cholinesterase inhibitor, insulin 총 9 종의 약물에 대한 체내 반응과 이와 관련된 유전적 변이 자료를 조사하여 catalog 로 제작하였다. 이는 환자의 유전적 요인을 고려하여 개개인에게 가장 적절한 약을 알맞은 용량으로 투여하는데 도움을 줄 수 있다. 조사한 9 종의 약물 중에서 Thiopurine 의 경우 6 명 중 2 명이 관련 snp 를 가지고 있으며 약물 투여 시 부작용으로 염증성 장 질환이 예측되었다. 그리고 insulin 의 경우 관련 snp 를 6 명 모두가 가지고 있어서 insulin 에 대한 저항성이 예측되었다. 이외의 나머지 약물에 대해서는 drug response 에 대한 유전자 변이를 가지고 있는 경우는 없었다.

Tamsulosin 은 방광 출구 근육 긴장을 완화하는 약물로, 배뇨장애 치료에 이용한다. 치료가 효과적이기 위해서는 tamsulosin 의 최대 혈청 농도(Cmax)는 약물이 독성을 일으키는 농도보다 낮거나 독성을 최소 이하로 유지할 정도여야 하고, 최저 혈청 농도는 약물이 효능을 보일 치료 범위 내에서 유지되어야 한다.⁸ GWAS 에서 조사한 결과 약물 반응과 관련된 snp 4 개를 확인할 수 있었다. rs16902947-A, rs7779057-C 와 같은 유전자 변이를 가지고 있으면 약물 복용 시 tamsulosin 의 최대 혈청 농도가 높게 나타나 치료 효과에 영향을 미친다. 6 명 모두 관련 유전자 변이는 발견되지 않았다.

Opiate 는 heroine, morphine 등 중추신경을 억제하는 마약성 진통제이다. 마약성 진통제의 가장 큰 부작용은 마약성 진통제 의존성[opiod dependence (OD)]로 심각한 사회적, 의학적 문제로 대두되고 있다. OD 위험성에 있어서 칼륨 및 칼슘 수송 및 신호 메커니즘은 필수적인 역할을 하는 것으로 보고되는데 GWAS 에서 조사한 결과 주로 칼륨 채널 sub family 단백질을 코딩하는 KCNC1, KCNG2 유전자의 변이인 snp 를 5 개 확인할 수 있었다. 관련 snp 를 가지고 있으면 opiate 투여 시 상대적으로 opiod dependence 의 위험성이 높게 나타난다고 보고된다. 6 명 모두 관련 유전자 변이는 발견되지 않았다.

Quetiapine 은 세로토닌, 도파민 작용을 억제하는 약물로 정신분열, 조울증 치료에 이용된다. Quetiapine 에 대한 반응과 FGF9, STOML2 와 같이 시냅스 기능, 신경 전달

⁸ Takata, Ryo, et al. "Impact of four loci on serum tamsulosin hydrochloride concentration." *Journal of human genetics* 58.1 (2013): 21-26.

물질 수용체 관련 유전자가 관련이 있다는 사실이 보고되었다. GWAS 에서 조사한 결과 약물 반응과 관련된 snp 18 개를 확인할 수 있었다. rs72790443, rs1471786 등의 snp 를 가지고 있으면 Quetiapine 의 치료 효과 정도에 영향을 미친다. 6 명 모두 관련 snp 는 나타나지 않았다.

Thiopurine 은 면역억제물질로 급성 림프성 백혈병, 자기면역장애 또는 장기이식 수혜자 치료에 이용된다. GWAS 에서 조사한 결과 약물 반응과 관련된 snp 24 개를 확인할 수 있었으며 주로 NUDT15, TPMT 와 관련된 유전자 변이인 경우가 많았다. 조사한 SNP 가 있으면 약물 부작용에 나타나는데 주로 염증성 장 질환(inflammatory bowel disease)이 발생한다. 6 명의 유전체 자료와 비교한 결과 2 명이 rs116855232 변이를 가지고 있었으며 thiopurine 투여 시 염증성 장 질환 발생 가능성이 높다는 예측이 나왔다.

Acetaminophen 은 cyclooxygenase(COX)의 작용을 억제하여 통증과 발열을 일으키는 prostaglandin 의 생성을 막는 약물로 해열 진통제로 이용된다. 체내로 들어온 대부분의 Acetaminophen 은 sulfation 되어 detoxification 되고 제거된다. 그러나 일부(5~7%)는 간에서 cytochromeP450(CYP450)에 의해 NAPQI 로 바뀌는데 rs6852435, rs2880961 등의 snp 를 가진 사람들은 NAPQI 가 conjugation 이 되지 않아 독성을 띄게 되어 간 독성(hepatotoxicity) 문제가 발생할 수 있다. GWAS 에서 조사한 결과 약물 반응과 관련된 snp 5 개를 확인할 수 있었으며 6 명 모두 관련 유전자 변이는 발견되지 않았다.

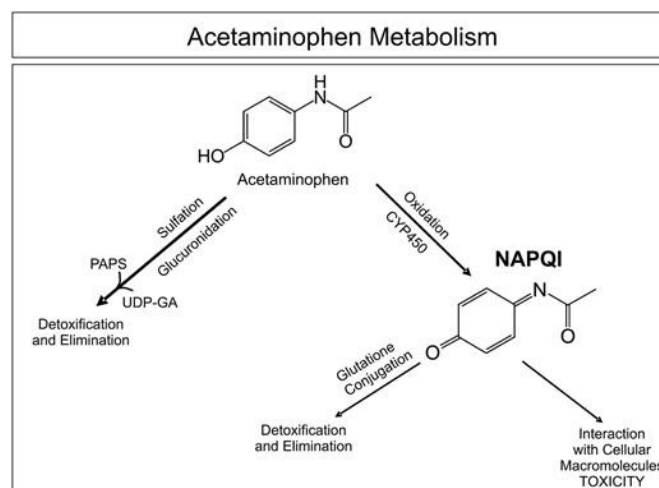


Figure 6. Metabolism of Acetaminophen

Wafarin 은 비타민 K 의 길항제로 미국에서 가장 널리 처방되는 경구용 항응고제이다. 하지만 요구되는 용량이 환자의 개인적 요인에 따라 차이가 크게 나타나는 약물 중 하나로 현재 미국에서 65 세 이상 노인의 약물 부작용과 관련된 입원의 33%를 차지하고 있다. CYP2C9, VKORC1 의 유전적 변이가 wafarin 복용량 변동성에 있어서 최대 30%를 설명한다고 보고되고 있다.⁹ 관련 유전자 변이를 가지고 있는 경우 wafarin 을 상대적으로 많이 복용해야 한다. GWAS 에서 조사한 결과 약물 반응과 관련된 snp 38 개를 확인할 수 있었으며 6 명 모두 관련 유전자 변이는 발견되지 않았다.

Methylphenidate 는 중추 신경에 작용하는 각성제로 ADHD 나 기면증 치료에 이용된다. 심혈관 질환과 관련된 유전자 변이를 가지고 있으면 해당 약물 처방 시 혈압 상승과 같은 심혈관계 관련 부작용이 나타날 수 있다고 보고된다.¹⁰ GWAS 에서 조사한 결과 관련 snp 7 개를 확인할 수 있었으며 6 명 모두 관련 유전자 변이는 발견되지 않았다.

Cholinesterase inhibitor 는 acetylcholine(Ach)를 분해하는 cholinesterase 를 억제하는 물질로 알츠하이머 치매에서 감소되는 acetylcholine 의 분해를 억제함으로써 신경 연결 내의 acetylcholine 농도를 증가시켜 인지기능 향상을 유도하게 된다.¹¹ GWAS 에서 조사한 결과 약물 반응과 관련된 snp 3 개를 확인할 수 있었으며 주로 단백질 kinase 관련 snp 임을 확인할 수 있었다. 관련 유전자 변이가 있는 경우 약물이 cholinesterase 에 대해 반응이 일어나지않아 치료효과가 미약하게 나타나는데 6 명 모두 관련 snp 는 발견되지 않았다.

Insulin 은 혈당을 낮추는 물질로 당뇨병 치료제로서 이용된다. 특정 snp 를 가지고 있으면 약물에 대한 저항성이 발생하여 치료 효과가 떨어지는 특징을 보인다. GWAS 에서 조사한 결과 약물 반응과 관련된 snp 25 개를 확인할 수 있었으며 6 명의 snp 자료와 비교한 결과 6 명 모두 rs1208-A 변이를 가지고 있었다.

⁹ Perera, Minoli A., et al. "Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study." *The Lancet* 382.9894 (2013): 790-796.

¹⁰ Mick, Eric, et al. "Genome-wide association study of blood pressure response to methylphenidate treatment of attention-deficit/hyperactivity disorder." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 35.2 (2011): 466-472.

¹¹ Martinelli-Boneschi, Filippo, et al. "Pharmacogenomics in Alzheimer's disease: a genome-wide association study of response to cholinesterase inhibitors." *Neurobiology of aging* 34.6 (2013): 1711-e7.

	A	B	C	D	E	F	SUM	TRAIT
rs1208	1	1	1	1	1	1	6	insulin
rs671	1	0	1	1	0	0	3	alcohol
rs1260326	0	1	0	0	1	0	2	caffeine
rs116855232	0	1	0	1	0	0	2	thiopurine
SUM	2	3	2	3	2	1	13	

Figure7. 6 명의 개인 유전체와 NCAD CATALOG 비교 분석 결과

결론

차세대염기서열분석(Next Generation Sequencing)이 도입되면서 유전체 분석은 비약적인 발전을 하였으며 환자의 다양한 유전정보를 정확하고 빠르게 얻음으로써 정확한 진단과 효과적인 처방이 가능해지고 있다. 또한 아직까지 밝혀지지 않은 질환의 원인과 약물의 메커니즘 규명에도 큰 도움을 받고 있다. 하지만 기하급수적으로 늘어나는 유전체 데이터 속에서 발견된 수많은 유전 변이의 임상적인 의미에 대한 해석이 큰 과제가 되고 있다.

GWAS 를 이용하면서 SNP 와 형질과의 인과관계가 구체적으로 명시되지 않아 유전체 분석을 진행하는데 불편함을 느꼈고 이를 보완할 필요성을 느꼈다. GWAS 의 다양한 유전정보 중 일상생활에서 특히 섭취 빈도가 높은 니코틴, 카페인, 술, 의약품의 metabolism 과 관련된 정보들을 선별하였고, 해당 SNP 에 따라 어떤 형질을 가질 수 있는지를 구체적으로 명시한 NCAD(Nicotine, Caffeine, Alcohol, Drugs)-metabolism catalog 를 만들었다. 그리고 건강한 성인 6 명의 상피세포를 채취하여 exome sequencing 을 진행한 유전체 자료와 제작한 NCAD catalog 를 이용해 분석하여 유전적 변이에 기반한 예측과 실제 개인의 특성이 일치하는지 확인하였다.

수많은 유전 변이의 임상적 의미에 대한 해석을 데이터베이스로 제작하는 작업은 유전체의 유전형 판단을 효율적이고 빠르게 수행하여 질병에 대한 개인별 차이를 알고 이에 맞춰 치료하는 정밀의학(Precision Medicine)에 한층 더 가까워질 수 있을 것이다. NGS 를 기반으로 임상유전체를 활용하는 정밀의학은 환자의 유전체를 확인하고 적절한 치료법을 선택함으로써 효율적인 치료를 기대할 수 있을 것이다.

<참고문헌>

- Benowitz, Neal L., Janne Hukkanen, and Peyton Jacob. "Nicotine chemistry, metabolism, kinetics and biomarkers." *Nicotine psychopharmacology*. Springer, Berlin, Heidelberg, 2009. 29–60.
- Choi JW Effects of ginsenosides on alcohol metabolism. Master thesis. Yeungnam University, 1983.
- Cornelis, Marilyn C., et al. "Genome-wide association study of caffeine metabolites provides new insights to caffeine metabolism and dietary caffeine-consumption behavior." *Human molecular genetics* 25.24 (2016): 5472–5482.
- Edwards, Alexis C., et al. "Meta-analysis of genetic influences on initial alcohol sensitivity." *Alcoholism: Clinical and Experimental Research* 42.12 (2018): 2349–2359.
- Gelernter, Joel, et al. "Genomewide association study of alcohol dependence and related traits in a Thai population." *Alcoholism: Clinical and Experimental Research* 42.5 (2018): 861–868.
- Hukkanen J, Jacob P, 3rd, Benowitz NL. Metabolism and disposition kinetics of nicotine. *Pharmacol Rev*. 2005c;57(1):79–115
- Linder MC. (1985). Nutritional biochemistry and metabolism. Elsevier, p. 48
- Mick, Eric, et al. "Genome-wide association study of blood pressure response to methylphenidate treatment of attention-deficit/hyperactivity disorder." *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 35.2 (2011): 466–472.
- Martinelli-Boneschi, Filippo, et al. "Pharmacogenomics in Alzheimer's disease: a genome-wide association study of response to cholinesterase inhibitors." *Neurobiology of aging* 34.6 (2013): 1711–e7.
- Patel, Yesha M., et al. "Novel association of genetic markers affecting CYP2A6 activity and lung cancer risk." *Cancer research* 76.19 (2016): 5768–5776.5768–5776
- Parra, Esteban J., et al. "Genome-wide association study of warfarin maintenance dose in a Brazilian sample." *Pharmacogenomics* 16.11 (2015): 1253–1263.
- Perera, Minoli A., et al. "Genetic variants associated with warfarin dose in African-American individuals: a genome-wide association study." *The Lancet* 382.9894 (2013): 790–796.
- Suhre, Karsten, et al. "Connecting genetic risk to disease end points through the human blood plasma proteome." *Nature communications* 8.1 (2017): 1–14.
- Takata, Ryo, et al. "Impact of four loci on serum tamsulosin hydrochloride concentration." *Journal of human genetics* 58.1 (2013): 21–26.

Yu, Hao, et al. "Five novel loci associated with antipsychotic treatment response in patients with schizophrenia: a genome-wide association study." *The Lancet Psychiatry* 5.4 (2018): 327-338.

권준택. "니코틴 대사와 CYP2A6 유전자 빈도에 관한 민족적 다형성." *Journal of Soonchunhyang Medical Science* 6.2 (2000): 285-292.

최지은. "페피노 추출액이 흰쥐의 알코올 대사에 미치는 영향." *국내석사학위논문 단국대학교*, 2008. 경기도

이승태, Next Generation Sequencing 기반 유전자 검사의 이해, *식품의약품안전처*, p.19

ACKNOWLEDGEMENT

2 개월 동안 이화여자대학교 생명정보학 연구실에서 함께 유전체 분석을 진행하고
NCAD metabolism catalog 제작한 김서영, 김선호, 류혜진, 백은하, 신지혜, 정소령
동기들에게 감사의 말씀을 전합니다