

SEPTEMBER 20XX

o

SBS Biohackathon 2023

Sheryl, Yi Da, Janelle

Problem Statement 2:

Exploring the science of obesity:
Using data analysis to determine the most prominent
factors that cause obesity and evaluate its efficacy in
treating obesity

SEPTEMBER 20XX

OI

Using the Obesity.csv dataset provided

Analysing Factors

O

Modifying the Dataset

FCVC	NCP	CAEC	SMOKE	CH2O	SCC	FAF	TUE	CALC	MTRANS
3	1	Frequently	no	1	no	0	0	Sometimes	Public_T
3	3.715118	Frequently	no	1.774576	no	0.10297	0.646423	no	Public_T
3	3.489918	Frequently	no	1.326694	no	0.791929	0.128394	no	Public_T
1.053534	3.378859	Sometimes	no	1	no	1.853425	0.861809	Sometimes	Public_T
3	3.263201	Sometimes	no	2.233274	no	1.557737	0.000355	Sometimes	Automot:
3	3.994588	Sometimes	no	2.548527	no	1.285976	0.267076	Sometimes	Automot:
2.530233	3.24934	Sometimes	no	2.387945	no	1.976341	1.672508	no	Automot:
3	3	Sometimes	no	2	no	0.854337	1	Sometimes	Public_T
2.8813	3.087544	Frequently	no	1.24818	no	1.952427	0.427461	no	Public_T
2.824559	3	Sometimes	no	2	no	0.533309	1.099764	Sometimes	Public_T
2.762325	1.163666	Sometimes	no	1.30491	no	0.25289	1.001405	Sometimes	Public_T
2.070964	1	Sometimes	no	1.676975	no	0	1.718513	Sometimes	Public_T
2.68601	1	Sometimes	no	1.310074	no	0	1.0647	Sometimes	Public_T
2.794197	3.409363	Sometimes	no	1.543021	no	2	1	Sometimes	Public_T
2.720701	3	Sometimes	no	2	no	1.862235	1	Sometimes	Public_T
2.880792	3.281391	Sometimes	no	1.960131	no	1.513029	1	Sometimes	Public_T
2.674431	3	Frequently	no	2.453384	no	1.321624	1.990617	no	Public_T
2	3.98525	Sometimes	no	2.654078	no	2.113151	2	no	Public_T
2.55996	3	Frequently	no	1.990788	no	0.486006	1.561272	no	Public_T
1.212908	3.207071	Sometimes	no	1.029703	no	2	1	Sometimes	Public_T
1.140615	3.471536	Sometimes	no	1	no	2	1	Sometimes	Public_T

Palechor, F. M., & De La Hoz Manotas, A. K. (2019). Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. Data in Brief, 25, 104344.
<https://doi.org/10.1016/j.dib.2019.104344>

Modifying the Dataset

```
df=pd.read_csv('/content/Obesity_dataset.csv')
for i in df['FCVC']:
    if i>=0 and i<1.5:
        df['FCVC']=df['FCVC'].replace(i,'Never')
    elif i>=1.5 and i<2.5:
        df['FCVC']=df['FCVC'].replace(i,'Sometimes')
    else:
        df['FCVC']=df['FCVC'].replace(i,'Always')

for i in df['FAF']:
    if i>=0 and i<0.5:
        df['FAF']=df['FAF'].replace(i,'0')
    elif i>=0.5 and i<1.5:
        df['FAF']=df['FAF'].replace(i,'1 to 2')
    elif i>=1.5 and i<2.5:
        df['FAF']=df['FAF'].replace(i,'2 to 4')
    else:
        df['FAF']=df['FAF'].replace(i,'4 to 5')
```

```
for i in df['CH2O']:
    if i>=0 and i<1.5:
        df['CH2O']=df['CH2O'].replace(i,'Less than 1L')
    elif i>=1.5 and i<2.5:
        df['CH2O']=df['CH2O'].replace(i,'Between 1-2L')
    else:
        df['CH2O']=df['CH2O'].replace(i,'More than 2L')

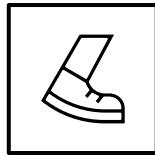
for i in df['NCP']:
    if i>=0 and i<1.5:
        df['NCP']=df['NCP'].replace(i,'1')
    elif i>=1.5 and i<2.5:
        df['NCP']=df['NCP'].replace(i,'2')
    elif i>=2.5 and i<3.5:
        df['NCP']=df['NCP'].replace(i,'3')
    else:
        df['NCP']=df['NCP'].replace(i,'4')

for i in df['TUE']:
    if i>=0 and i<0.5:
        df['TUE']=df['TUE'].replace(i,'0 to 2')
    elif i>=0.5 and i<1.5:
        df['TUE']=df['TUE'].replace(i,'3 to 5')
    else:
        df['TUE']=df['TUE'].replace(i,'>5')
```

Modifying the Dataset

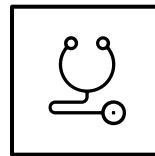
FCVC	NCP	CAEC	SMOKE	CH2O	SCC	FAF	TUE	CALC
Always		4 Sometimes	no	Less than 1L	no	2 to 4	3 to 5	Sometimes
Never		4 Frequently	no	Less than 1L	no	1 to 2	3 to 5	Sometimes
Never		4 Frequently	no	Between 1-2L	no	2 to 4	3 to 5	Sometimes
Never		4 Frequently	no	Less than 1L	no	2 to 4	3 to 5	Sometimes
Always		3 Frequently	no	Less than 1L	no	2 to 4	0 to 2	Sometimes
Always		1 Frequently	no	Less than 1L	no	1 to 2	0 to 2	Sometimes
Always		1 Frequently	no	Between 1-2L	no		0 3 to 5	Sometimes
Sometimes		3 Sometimes	no	Between 1-2L	no	4 to 5	3 to 5	no
Sometimes		3 Sometimes	no	Between 1-2L	no	2 to 4	3 to 5	no
Always		3 Sometimes	no	Between 1-2L	no	2 to 4	3 to 5	no
Sometimes		4 Frequently	no	Less than 1L	no	1 to 2	0 to 2	no
Always		3 Frequently	no	More than 2L	no	1 to 2	3 to 5	no
Sometimes		3 Frequently	no	Less than 1L	no	2 to 4	>5	no
Sometimes		3 Sometimes	no	Less than 1L	no		0 3 to 5	Sometimes
Never		3 Sometimes	no	Between 1-2L	no		0 3 to 5	Sometimes
Sometimes		4 Sometimes	no	Less than 1L	no	1 to 2	3 to 5	Sometimes
Never		2 Sometimes	no	Less than 1L	no	1 to 2	>5	Sometimes
Sometimes		3 Sometimes	no	Between 1-2L	no	2 to 4	3 to 5	Sometimes
Always		3 Sometimes	no	Less than 1L	no		0 0 to 2	Sometimes

DATA CATEGORIES



ORDINAL

CAEC
CALC
FCVC
CH2O
FAF
TUE
NObeyesdad



CONTINUOUS

Height
Weight
Age

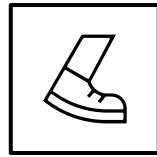


NOMINAL

Gender
Family History
Smoke
FAVC
SCC
NCP
MTRANS

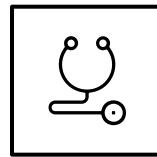
Dependent variable!

DATA CATEGORIES



ORDINAL

Spearman's Coefficient



CONTINUOUS

Spearman's Coefficient



NOMINAL

2 groups : Mann-Whitney U
>2 groups : Kruskal Wallis

Top 3 Categorical Variables that affect Obesity

1. CAEC = Food consumption between meals
2. Family History
3. FAVC = Frequent High Caloric Food Consumption
4. SCC = Calories consumption monitoring
5. MTRANS = Type of transportation used

Arranged in ascending p-value

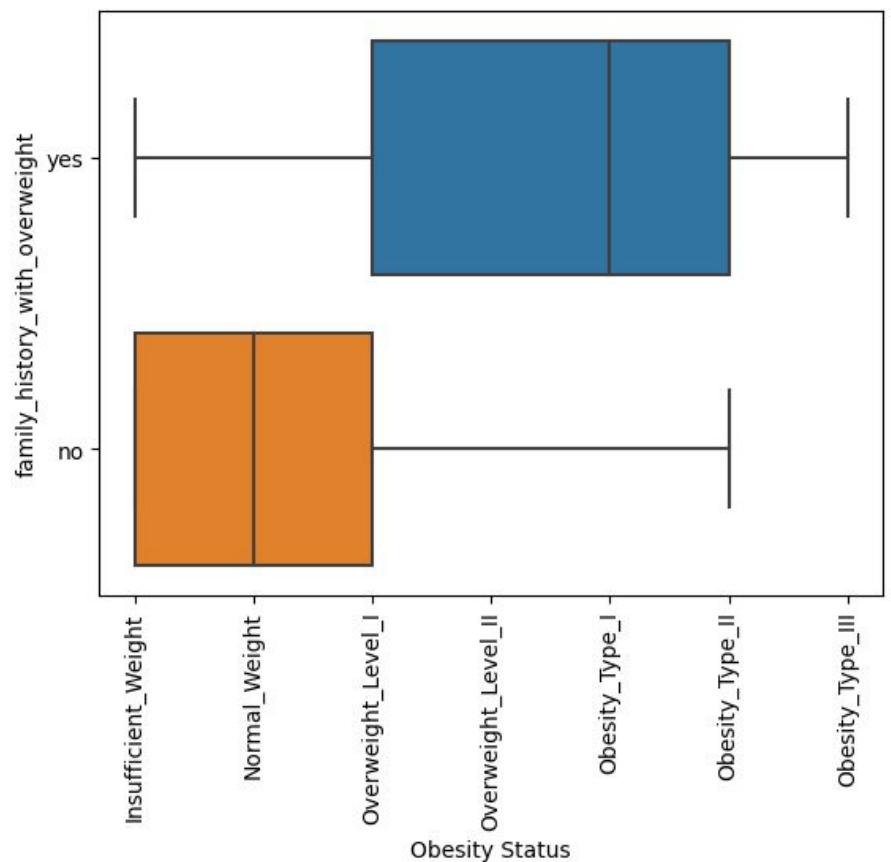
Variable	Chi-Sq	P-Value
NObeyesdad	4222.000000	0.000000e+00
CAEC	609.161875	2.462673e-128
family_history_with_overweight	570.042440	1.647619e-124
FAVC	176.034488	5.951093e-39
SCC	77.843730	1.248673e-17
MTRANS	85.708854	3.444571e-15
CALC	70.547623	3.156927e-13
Weight	3627.426276	1.220522e-12
NCP	1551.364154	6.890582e-08
Gender	20.492198	3.549569e-05
Age	2087.694159	1.066510e-04
FCVC	1719.014072	3.991022e-02
FAF	2477.894808	7.516805e-02
Height	3249.289518	9.739292e-02
SMOKE	2.048749	3.590210e-01
TUE	2232.087374	6.358267e-01
CH20	2479.578742	7.765557e-01

O

NOMINAL (Mann Whitney U)

NOMINAL (2 GROUPS) (MANNWHITNEYU):

	Variable	Shorthand	U	p value
3	Smoke	SMOKE	44883	0.001431
1	Family History	family_history_with_overweight	86345.5	7.48672e-117
2	High Caloric Food Frequency	FAVC	126670	1.62721e-30
4	Calories Consumption Monitoring	SCC	148300	4.13225e-19
0	Gender	Gender	580517	0.0891216



Family History 86345.5 7.48672e-117

THEORIES

- 1) Obesity is genetically predisposed
- 2) Family environment may influence lifestyle habits that increase one's susceptibility to developing obesity

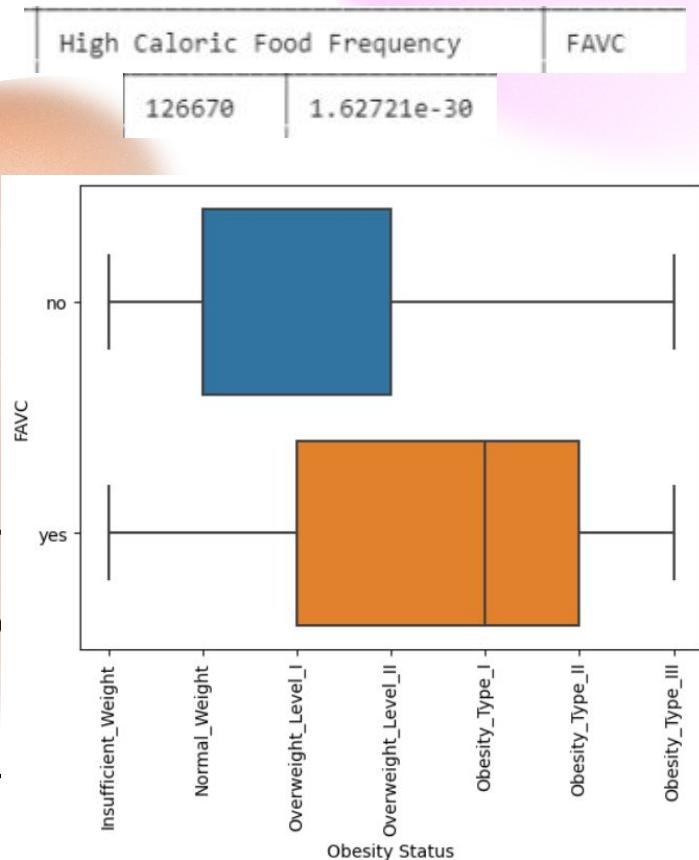
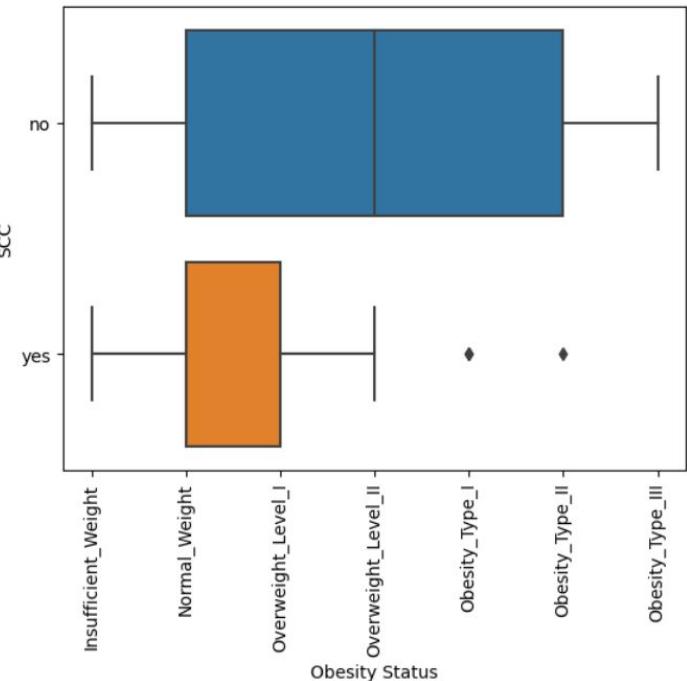
O

NOMINAL (Mann Whitney U)

O

NOMINAL (Mann Whitney U)

Calories Consumption Monitoring		SCC
148300	1.62721e-30	4.13225e-19

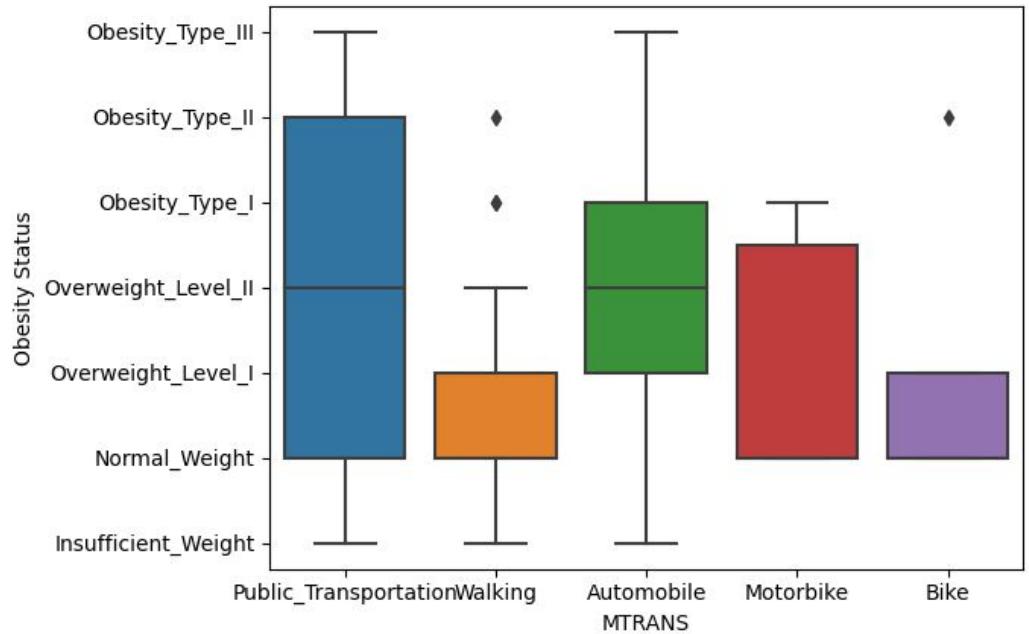


O

NOMINAL (Kruskal-Wallis)

NOMINAL (>2 GROUPS) (KRUSKAL-WALLIS):

	Variable	Shorthand	H	p value
1	Transportation Used	MTRANS	79.6752	3.60338e-17
0	Number of main meals	NCP	17.869	0.00130901



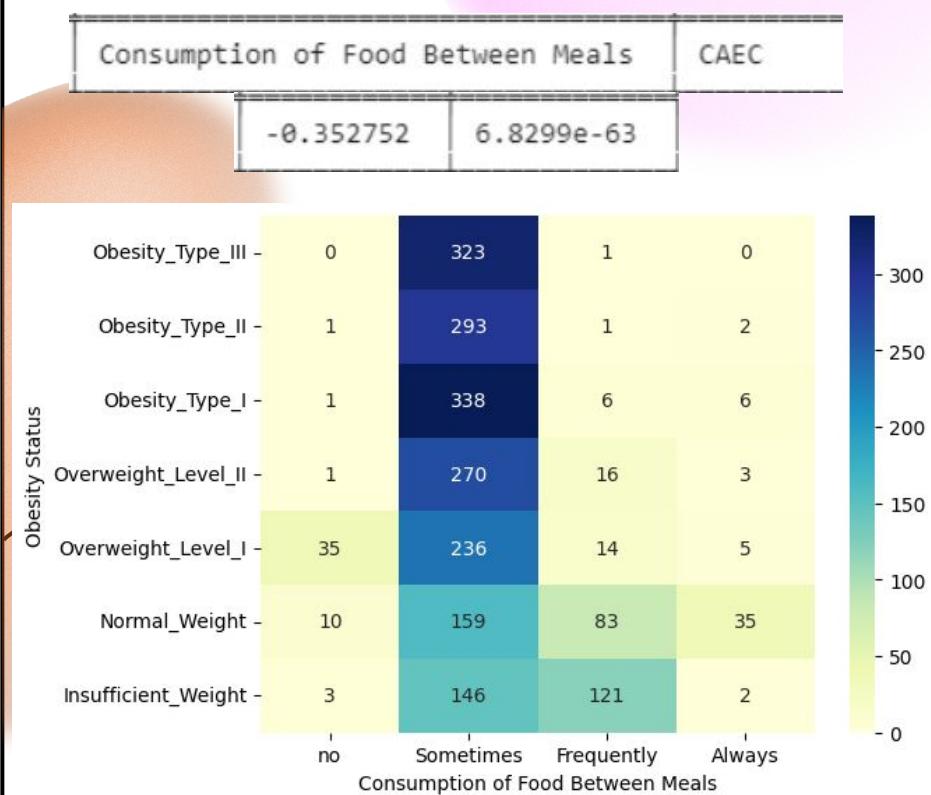
THEORIES

Depends on how physically demanding the transport type is

NOMINAL (Kruskal-Wallis)

ORDINAL (SPEARMAN):

	Variable	Shorthand	corr	p value
1	Consumption of Food Between Meals	CAEC	-0.352752	6.8299e-63
0	Frequency of Vegetable Consumption	FCVC	0.207816	5.00742e-22
3	Physical Activity Frequency	FAF	-0.19027	1.16993e-18
5	Consumption of Alcohol	CALC	0.168379	6.84244e-15
2	Consumption of Water Daily	CH2O	0.137553	2.20418e-10
4	Time Using Technology Devices	TUE	0.0598232	0.00596956



LIMITATIONS

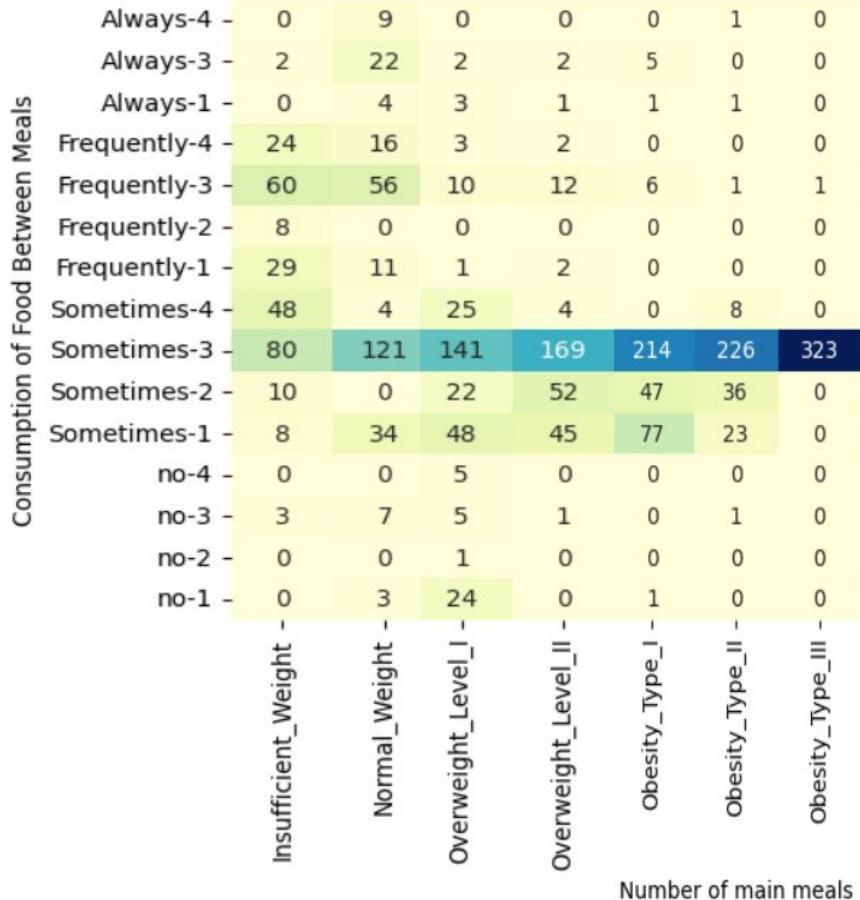
- 1) Fails to consider type of food consumed
- 2) Fails to consider portion size

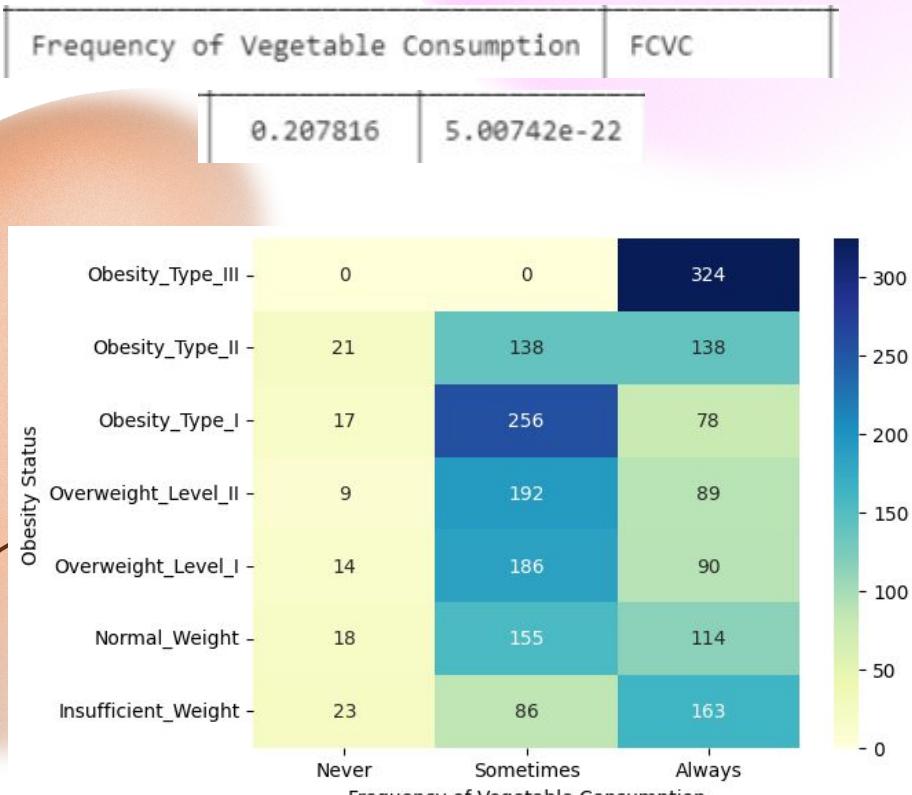
THEORIES

- 1) More consumption between meals = more time to digest = less need to store as fat = lower risk of obesity
- 2) Consistency is key! Consistent eating habits = lower risk of obesity. Needs more research

O

NOMINAL (kruskal-Wallis)





LIMITATIONS

- 1) Fails to consider type of vegetable consumed and its nutrient content
- 2) Fails to consider portion size

THEORIES

- 1) People who eat more vegetables = people who eat more in general? Therefore, not vegetables in particular that play a role
- 2) Certain components in vegetables may have underlying effects on the metabolism of our body to result in obesity. More research needed.
- 3) It may be that people who are obese/at risk of obesity increase their vegetable intake in hopes that a healthier diet can help them lose weight

O

ORDINAL (Spearman)

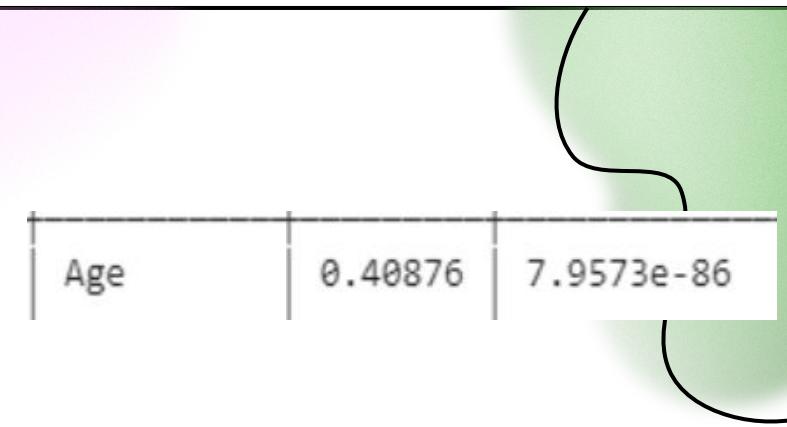
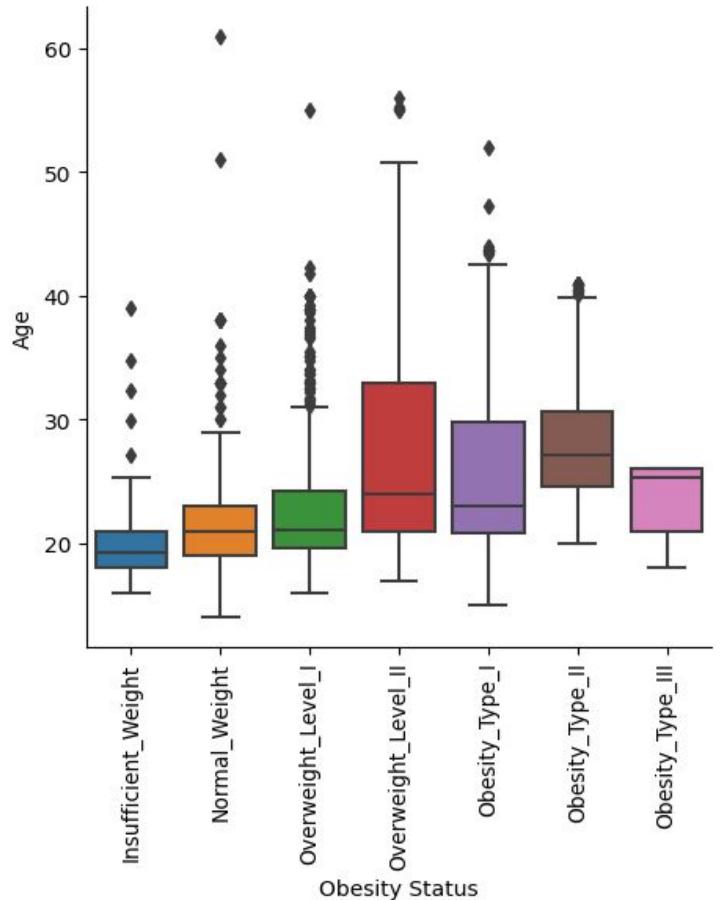
Physical Activity Frequency	FAF	-0.19027	1.16993e-18
Consumption of Alcohol	CALC	0.168379	6.84244e-15
Consumption of Water Daily	CH2O	0.137553	2.20418e-10

O

CONTINUOUS (Spearman)

CONTINUOUS (SPEARMAN):

	Variable	Shorthand	corr	p value
2	Weight	Weight	0.92132	0
0	Age	Age	0.40876	7.9573e-86
1	Height	Height	0.12683	4.99108e-09



THEORIES

- 1) Older people may lead more sedentary lifestyles
- 2) Older people have slower metabolism

CONTINUOUS (Spearman)

Top Factors from all Tests

Genetic Biomarkers

Age & Family History

Eating Habits

CAEC, FAVC, SCC

Physical Activity Related

MTRANS

O

02

Significant Factor: Genetic Predisposition

GEO help: Mouse over screen elements for information.

Scope: Format: Amount: GEO accession: **Series GSE9624**[Query DataSets for GSE9624](#)

Status	Public on Oct 01, 2008
Title	Differential gene expression in omental adipose tissue from obese children
Organism	<i>Homo sapiens</i>
Experiment type	Expression profiling by array
Summary	<p>Characterization of genes associated with adipose tissue is key to understanding the pathogenesis of obesity and developing treatments for this disorder. Differential gene expression in the adipose tissue has been described in adulthood but none studies have been developed on childhood. The purpose of this study was to compare gene expression in omental adipose tissue from obese prepubertal and normal weight children. We selected 5 obese (BMI adjusted for age and sex z score >2) and 6 normal weight children. RNA was extracted from omental adipose tissue biopsies and cRNA was hybridized on the human genome U133 Plus 2.0 Arrays (Affymetrix®). Microarray experiments were performed for each sample, and selected group of gene expression values were confirmed with real-time RT-PCR in 10 obese and 10 normal weight prepubertal children. 1276 genes were found to be differentially expressed at P<0.05. Of those differential genes, 201 were upregulated ($F_c>2$) and 42 were downregulated ($F_c<-2$). Genes involved in metabolic and signalling pathways were altered in childhood obesity.</p> <p>Keywords: disease state analysis</p>
Overall design	Adipose tissue and blood samples were obtained from 27 children, 14 obese (BMI adjusted for age and sex z score > 2) and 13 non obese undergoing appendix surgery. About 400 mg of adipose tissue was taken and immediately immersed in RNAlater solution and stored at -80°C for gene expression analysis. Informed consent was obtained from all patients after the nature of the study was explained, and the experimental design was approved, from an ethical and scientific standpoint, by the Hospital's Ethical Committee responsible for research.
Contributor(s)	Aguilera CM , Tofe I , Suarez A , Gomez-Llorente C , Cañete R , Gil A
Citation(s)	Aguilera CM, Gomez-Llorente C, Tofe I, Gil-Campos M et al. Genome-wide expression in visceral adipose tissue from obese prepubertal children. <i>Int J Mol Sci</i> 2015 Apr 8;16(4):7723-37. PMID: 25856673

Gene Expression Dataset from adipose tissue of normal weight and obese prepubertal children

Children have yet to have long exposure to lifestyle environmental factors that may influence obesity level

Study done in Spain versus previous dataset study conducted with Mexican, Columbian and Peru individuals whom may have Spanish descent

	A	B	C	D	E	F	G	H	I	J	K	L	M
41	ISample_type	RNA	RNA	RNA	RNA	RNA	RNA	RNA	RNA	RNA	RNA	RNA	
42	ISample_channel_count	1	1	1	1	1	1	1	1	1	1	1	
43	ISample_source_name_ch1	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose	Omental adipose tissue, o	
44	ISample_organism_ch1	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	
45	ISample_characteristics_ch1	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental	tissue: omental adipose
46	ISample_characteristics_ch1	group: normal	group: normal	group: normal	group: normal	group: normal	group: normal	group: obese	group: obese	group: obese	group: obese	group: obese	group: obese
47	ISample_characteristics_ch1	gender: M	gender: M	gender: F	gender: M	gender: M	gender: M	gender: M	gender: M	gender: M	gender: M	gender: M	gender: M
48	ISample_characteristics_ch1	age: 10 year	age: 7 year	age: 8 year	age: 10 year	age: 7 year	age: 7 year	age: 11 years	age: 9 years	age: 12 years	age: 9 years	age: 10 years	
49	ISample_characteristics_ch1	z score bmi: -1.4	z score bmi: 0.4	z score bmi: 0.12	z score bmi: -1.8	z score bmi: -0.6	z score bmi: -0.3	z score bmi: 2	z score bmi: 4.2	z score bmi: 2.3	z score bmi: 4.2	z score bmi: 2.72	
76	ID_REF	GSM243215	GSM243216	GSM243217	GSM243218	GSM243219	GSM243220	GSM243225	GSM243226	GSM243227	GSM243228	GSM243275	
77	1007_s_at	0.99751663	0.96097976	1.5767715	1.0237662	0.8849869	0.99067724	1.9664539	0.97925806	1.7902808	2.5839422	1	
78	1053_at	1.0533872	1.0356168	0.9108959	0.6967738	1.4531766	1.1084821	0.4729472	0.44015634	0.47881413	1.1809167	1	
79	117_at	1.3381977	0.87183416	1	1.4466145	1.6374854	1.656227	0.43038094	0.58551186	0.60171825	0.91075563	3.2734404	
80	121_at	0.6091248	0.97019064	1.3350986	0.9151971	1.1050855	1.3639722	1	1.0357399	0.8812516	0.88187283	1.3702561	
81	1255_g_at	0.8959755	0.7171478	0.16077964	0.57525766	0.1222709	0.81999993	0.30999997	0.7146315	0.24734041	0.10043999	0.3871292	
82	1294_at	0.7097115	0.7447934	1.0231184	0.6740467	1.0715216	0.89304835	1	1.1371415	1.0840693	0.6969424	1.4770888	
83	1316_at	0.912628	0.62323946	1	1.316553	1.3920423	0.60145235	1.291735	0.84689	2.2063186	1.3608313	0.8358914	
84	1320_at	1.2343444	0.14227068	0.18262091	1.0897492	1	1.4613689	0.09851414	0.7432263	0.12157065	1.3762145	1.0662154	
85	1405_i_at	1.6260741	0.72823423	0.8616563	0.29577723	0.7746381	0.5714997	1	1.074396	1.5173014	1.3164555	1.6991435	
86	1431_at	0.4617886	1	1.2194351	0.84581906	1.357769	1.2021165	1.5233716	3.1128488	0.7577171	0.55177146	0.98354983	
87	1438_at	0.4972239	1.6341035	0.99999994	1.594795	1.9349432	2.5402136	0.20675635	0.3976252	0.12757309	1.0145123	0.15491888	
88	1487_at	1.721643	0.90215194	0.9282138	1.2428621	0.94474745	1.0603511	0.73985386	0.5900059	1	1.6815166	1.4455881	
89	1494_f_at	0.74874604	1.2893718	1	1.0339197	0.6782819	0.78380495	1.1218333	1.3389193	0.96022546	1.0126865	0.87778807	
90	1552256_a_at	1.0248879	1	0.87998676	0.9869518	1.1381584	0.77410114	1.5877707	1.7767727	1.5349771	0.7351664	0.87500703	
91	1552257_a_at	1.8980603	1	0.85355395	1.302797	1.2098755	0.9431133	0.76654094	0.7218677	1.0647159	0.96783626	1.1213201	
92	1552258_a_at	0.6072004	1	0.65521000	1.5027600	0.20052004	1.4021405	1.4400275	1.0000240	0.5077307	0.55520010	1.7112000	

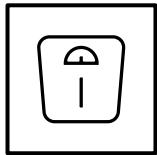
ID_REF = gene identifier searchable on BioGPS

GSM##### = child sample gene expression values

Normal Healthy Weight

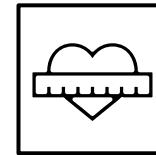
Obese

Statistical Concepts Used



Standard deviation

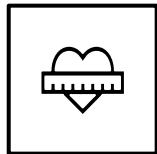
To identify whether within each group of obese vs healthy samples gene expression is similar



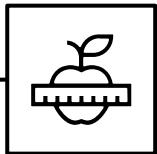
t-test

To identify whether there is significant difference between gene expression in obese vs healthy group

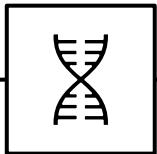
METHODOLOGY



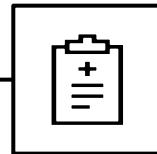
Find std within healthy vs obese sample groups for expression of each gene



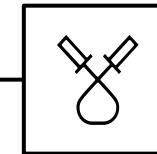
Compare mean expression between healthy vs obese for each gene using t-test



Determine genes with significant difference in expression between healthy vs obese



Narrow down top 50 genes of smallest p-value



Determine gene correlation to obesity

```
def std_dev(data):
    mean = np.mean(data)
    std = np.sqrt(np.sum((data - mean)**2) / (len(data) - 1))
    return std

def mean(data):
    return np.mean(data)

df = pd.read_csv('GSE9624_series_matrix.txt', sep='\t', skiprows=75,
index_col='ID_REF', skipfooter=1, engine='python')
|
healthy = df.iloc[:, 1:7]
obese = df.iloc[:, 7:12]
healthy_std = healthy.apply(std_dev, axis=1)
obese_std = obese.apply(std_dev, axis=1)
healthy_mean = healthy.apply(mean, axis=1)
obese_mean = obese.apply(mean, axis=1)
```

```
from scipy.stats import ttest_ind
pvalues = []
for i in range(len(healthy.index)):
    t, p = ttest_ind(healthy.iloc[i,:].tolist(), obese.iloc[i,:].tolist())
    pvalues.append(p)

updown = []
for i in range(len(healthy.index)):
    if healthy_mean[i] > obese_mean[i]:
        updown.append("Down")
    elif healthy_mean[i] < obese_mean[i]:
        updown.append("Up")
    else:
        updown.append("-")

result = pd.DataFrame({'ID_REF': df.index, 'healthy_mean': healthy_mean,
'healthy_std_dev': healthy_std, 'obese_mean': obese_mean, 'obese_std_dev':
obese_std, 't_test_p_value': pvalues,
'up_or_downregulated_in_obese':updown})
print(result)
```

ID REF	SYMBOL	UP/DOWN REGULATED	NAME
1562659_at	LINC02139	Up	long intergenic non-protein coding RNA 2139
234666_at		Up	unknown
238364_x_at	GLI4	Up	GLI family zinc finger 4
1555262_a_at	MAGH1	Up	membrane associated guanylate kinase, WW and PDZ domain containing 1
224253_at	EXOC5	Up	exocyst complex component 5
200748_s_at	FTH1	Up	ferritin heavy chain 1
215584_at	HECW1	Down	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1
238138_at	CANX	Up	calnexin
241674_s_at		Up	unknown
210496_at		Down	unknown
221012_s_at	TRIM8	Up	tripartite motif containing 8
221865_at	TMEM268	Up	transmembrane protein 268
236616_at	PSD4	Down	pleckstrin and Sec7 domain containing 4
239656_at	LHFPL3-AS2	Up	LHFPL3 antisense RNA 2
235735_at	TNFSF8	Up	TNF superfamily member 8
213389_at	ZNF592	Up	zinc finger protein 592
1555856_s_at	AKR1C1	Up	aldo-keto reductase family 1 member C1
235418_at	FAHD2A	Up	fumarylacetoacetate hydrolase domain containing 2A
236104_at	HNRNPLL	Up	heterogeneous nuclear ribonucleoprotein L like
1559551_at		Down	unknown
212307_s_at	OGT	Down	O-linked N-acetylglucosamine (GlcNAc) transferase
222170_at	ZNF213-AS1	Down	ZNF213 antisense RNA 1 (head to head)
227830_at	GABRB3	Up	gamma-aminobutyric acid type A receptor subunit beta3
201366_at	ANXA7	Up	annexin A7
211054_at	INVS	Up	inversin
57082_at	LDLRAP1	Up	low density lipoprotein receptor adaptor protein 1
1563546_at		Up	unknown
1562371_s_at	VWA3B	Up	von Willebrand factor A domain containing 3B
221367_at	MOS	Down	MOS proto-oncogene, serine/threonine kinase
238569_at	GABBR1	Up	gamma-aminobutyric acid type B receptor subunit 1
1552359_at	MCMDC2	Down	minichromosome maintenance domain containing 2
233491_at	RPL26L1-AS1	Down	RPL26L1 antisense RNA 1
234059_at		Down	unknown
241150_at	SPTAN1	Up	spectrin alpha, non-erythrocytic 1
1569790_at		Up	unknown
201471_s_at	SQSTM1	Up	sequestosome 1
227673_at	POLR1H	Up	RNA polymerase I subunit H
210390_s_at	CCL15	Down	C-C motif chemokine ligand 15
214884_at	MCF2	Up	MCF2 cell line derived transforming sequence
243530_at		Up	unknown
233114_at		Down	unknown
240406_at		Up	unknown
241317_at		Down	unknown
217614_at	ARHGAP45	Up	Rho GTPase activating protein 45
1555199_at	GOSR1	Down	golgi SNAP receptor complex member 1
215723_s_at	PLD1	Down	phospholipase D1
213818_x_at	TMEM223	Up	transmembrane protein 223
204106_at	TESK1	Up	testis associated actin remodelling kinase 1
1557336_at	SUCLG2-DT	Down	SUCLG2 divergent transcript
213138_at	ARID5A	Up	AT-rich interaction domain 5A

Genes arranged in ascending t-test p-value with regards to gene expression between obese and healthy samples

Gene Function Types	Count (include overlaps)
Cell Adhesion, Extracellular Matrix and Cytoskeleton Proteins	4
Transcription Regulation	9
Cell proliferation, differentiation, death	7
Cell Signalling	15
Immune Response	4
Transport	5
Metabolism	5
Protein Synthesis	3
Unknown	11

UPREGULATED

O

ID REF	SYMBOL	UP/DOWN REGULATED	NAME	Description
LINC02139				(lncRNA) genes have diverse features that distinguish them from mRNA-encoding genes and exercise functions such as remodelling chromatin and genome architecture, RNA stabilization and transcription regulation, including enhancer-associated activity.
1562659_at		Up	long intergenic non-protein coding RNA 2139	
234666_at		Up	unknown	
GLI4				Predicted to enable DNA-binding transcription factor activity, RNA polymerase II-specific and RNA polymerase II cis-regulatory region sequence-specific DNA binding activity. Predicted to be involved in regulation of transcription, DNA-templated. Predicted to be located in nucleus
238364_x_at	MAGI1	Up	GLI family zinc finger 4	membrane associated guanylate kinase, WW and PDZ domain containing 1
1555262_a_at	EXOC5	Up		ATP, protein and alpha actinin binding protein
224253_at	exocyst complex component 5	Up		protein component of exocyst transporter complex
200748_s_at	FTH1	Up	ferritin heavy chain 1	Encodes ferritin heavy chain, a major non-haem iron storage protein Calcium-binding molecular chaperone protein that interacts with newly synthesized N-linked glycoproteins in the endoplasmic reticulum.
238138_at	CANX	Up	calnexin	
241674_s_at		Up	unknown	
221012_s_at	TRIM8	Up	tripartite motif containing 8	Encodes a member of the tripartite motif (TRIM) protein family. Suspected to be an E3 ubiquitin-protein ligase
TMEM268				transmembrane protein 268
221865_at		Up		Family of protein with unknown function Predicted to be integral component of membrane.
239656_at	LHPFL3-AS2	Up	long noncoding RNA tumour suppressor gene	LHPFL3 antisense RNA 2
TNFSF8				Encodes a cytokine that belongs to the tumor necrosis factor (TNF) ligand family. Inhibitory role in modulating Ig class switch in B cells Induce cell death and regulates cell proliferation in lymphoma cell lines
235735_at		Up		
ZNF592				zinc finger protein 592
213389_at	AKR1C1	Up	aldo-keto reductase family 1 member C1	Gene is thought to play a role in a complex developmental pathway and the regulation of genes involved in cerebellar development
1555856_s_at		Up		metabolic oxidoreductase enzyme involved in lipid metabolism involved in cholesterol homeostasis
FAHD2A				fumarylacetoacetate hydrolase domain containing 2A Predicted to enable hydro-lyase activity.
235418_at	HNRNPLL	Up		Fumarylacetoacetate hydrolase catalyses the hydrolytic cleavage of a carbon-carbon bond in fumarylacetoacetate to yield fumarate and acetoacetate as the final step in phenylalanine and tyrosine degradation
236104_at	GABRB3	Up	heterogeneous nuclear ribonucleoprotein L like	A master regulator of activation-induced alternative splicing in T cells After splicing of CD45, a tyrosine phosphatase essential for T-cell development and activation
227830_at				Encoded protein is one the subunits of a multi-subunit chloride channel that serves as the receptor for gamma-aminobutyric acid, a major inhibitory neurotransmitter of the mammalian nervous system
201366_at	ANXA7	Up	annexin A7	A member of the annexin family of calcium-dependent phospholipid binding proteins.
211054_at	INVS	Up	inversin	calmodulin binding renal developmental protein involved in Wnt signaling pathway
57082_at	LDLRAP1	Up	low density lipoprotein receptor adaptor protein 1	Cytosolic protein with a phosphotyrosine binding protein that interacts with the cytoplasmic tail of the LDL receptor
1563545_at		Up	unknown	
1562371_s_at	VWA3B	Up	von Willebrand factor A domain containing 3B	Expression in fallopian tube, respiratory epithelia and testis.
238569_at	GABBR1	Up	gamma-aminobutyric acid type B receptor subunit 1	receptor for Gamma-aminobutyric acid (GABA)
241150_at	SPTAN1	Up	spectrin alpha, non-erythrocytic 1	calmodulin binding actin capping protein
1563790_at		Up	unknown	
201471_s_at	SQSTM1	Up	sequestosome 1	Mainly protein binding
227673_at	POLR1H	Up	RNA polymerase I subunit H	RNA polymerase 1 transcribes rRNA
214884_at	MCF2	Up	MCF2 cell line derived transforming sequence	guanyl-nucleotide exchange factor activity
243530_at		Up	unknown	
240406_at		Up	unknown	
217614_at	ARHGAP45	Up	Rho GTPase activating protein 45	GTPase activator
213818_x_at	TMEM223	Up	transmembrane protein 223	Mitochondrial ribosome-associated protein
204106_at	TESK1	Up	testis associated actin remodelling kinase 1	serine/threonine protein kinase
213138_at	ARID5A	Up	AT-rich interaction domain 5A	transcription factor that binds to AT rich region

- LDLRAP1 show upregulation of lipid metabolism
- AKR1C1 show upregulation of lipid transport and storage

DOWN\REGULATED

ID REF	SYMBOL	UP/DOWN REGULATED	NAME	Description
	HECW1			Predicted to enable ubiquitin protein ligase activity. Predicted to be involved in several processes, including cellular protein metabolic process; negative regulation of sodium ion transmembrane transporter activity; and regulation of dendrite morphogenesis.
215584_at		Down	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1	Located in cytosol.
210496_at		Down	unknown	
236616_at	PSD4	Down	pleckstrin and Sec7 domain containing 4	lipid binding guanine nucleotide exchange factor
1559551_at		Down	unknown	
212307_s_at	OGT	Down	O-linked N-acetylglucosamine (GlcNAc) transferase	Encodes a glycosyltransferase that catalyzes the addition of a single N-acetylglucosamine in O-glycosidic linkage to serine or threonine residues
222170_at	ZNF213-AS1	Down	ZNF213 antisense RNA 1 (head to head)	long non coding RNA for
221367_at	MOS	Down	MOS proto-oncogene, serine/threonine kinase	proto oncogene for Serine/threonine-protein kinase
1552359_at	MCMDC2	Down	minichromosome maintenance domain containing 2	protein localised to actin
233491_at	RPL26L1-AS1	Down	RPL26L1 antisense RNA 1	long non coding RNA that may encode ribosomal protein
234059_at		Down	unknown	
210390_s_at	CCL15	Down	C-C motif chemokine ligand 15	cytokine involved in chemotaxis process
233114_at		Down	unknown	
241317_at		Down	unknown	
1555199_at	GOSR1	Down	golgi SNAP receptor complex member 1	essential to SNARE
	PLD1		phospholipase D1	phospholipase selective for phosphatidylcholine involved in signal transduction, membrane trafficking, and the regulation of mitosis
215723_s_at		Down		
1557336_at	SUCLG2-DT	Down	SUCLG2 divergent transcript	transcripte variant encodes a GTP-specific beta subunit of succinyl-CoA synthetase with ligase activity

- SUCLG2-DT show downregulation of citric acid cycle
- OGT downregulation leads to cell signalling promoting basal lipolysis which reduce sensitivity to nutrient-induced lipolysis

O

SEPTEMBER 20XX

THANK YOU