INVESTIGATION OF THE EFFECTS OF NICOTINE AND LEVAMISOLE ON SW620 COLON ADENOCARCINOMA CELLS USING A CUSTOMIZED R-ROUTINE FOR AUTOMATED MICROARRAY ANALYSIS

A THESIS SUBMITTED TO

THE DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS AND

INSTITUTE OF ENGINEERING AND SCIENCE OF

BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

By Muammer ÜÇAL

August 2010

To my family
and to my daisy, Saniye

ABSTRACT

INVESTIGATION OF THE EFFECTS OF NICOTINE AND LEVAMISOLE ON SW620 COLON ADENOCARCINOMA CELLS USING A CUSTOMIZED R-ROUTINE FOR AUTOMATED MICROARRAY ANALYSIS

Muammer ÜÇAL

MSc. in Molecular Biology and Genetics

Supervisor: Assoc. Prof. Dr. Özlen KONU

August 2010

Nicotine, the addictive component of tobacco, shows proliferative and antiapoptotic activity in cancer cells. Levamisole, an antihelmintic, on the other hand, has been tested as an additive chemotherapeutic agent and in treatment of nephrotic syndrome. Nicotine and levamisole are both agonists of nicotinic acetylcholine receptors; effects of these two agents have not been studied in colon cancer transcriptome. In this study, nicotine and levamisole exposed SW620 colon cancer cells, at a dose of 1 μM for 7 days, were studied with respect to changes in expression using microarrays. For data analysis, a custom R-routine which makes extensive use of open source R-BioConductor Project and associated packages has been written; and it is composed of three modules: *QualCont* module performs quality controls supported with several visual

representations and normalization of the raw data; *DEGidentifier* module performs identification of differentially expressed genes (DEGs) supported with Heatmap and Venn Diagram; and finally *FuncAnn* module, composed of four main functions, performs functional annotation and analysis of the results in terms of Gene Ontology, pathway analysis, supported with graphics and pathway networks

All arrays passed the quality control criteria. In this study, we show that nicotine and levamisole treatments affect the transcriptome similarly in SW620 cells in terms of both the differential expression and functional analysis. Our findings implicate nicotine and levamisole in steroid biosynthesis, cholesterol biosynthesis and aminoacid degradation metabolisms. Confirmatory analyses have been performed by real time RT-PCR for a selected set of genes.

Keywords: nicotine, levamisole, SW620, colon cancer, custom R routine, automated microarray analysis

ÖZ

NİKOTİN VE LEVAMİSOLÜN SW620 KOLON ADENOKARSİNOM HÜCRELERİ ÜZERİNDEKİ ETKİLERİNİN ARAŞTIRILMASI VE OTOMATİK MİKRODİZİ ANALİZİ İÇİN ÖZEL R PROGRAM RUTİNİ GELİŞTİRİLMESİ VE UYGULANMASI

Muammer ÜÇAL

Moleküler Biyoloji ve Genetik Yüksek Lisansı

Tez danışmanı: Doç. Dr. Özlen KONU

August 2010

Sigaranın bağımlılık yapıcı madde olarak bileşeni nikotin, hücre çoğalımına neden olmak suretiyle kanserojen etki gösterir. Levamisol bir antihelmintik olarak kullanımdaysa da kolon kanserinde bağışıklığı arttırmak amaçlı kemoterapi ajanı olarak ve steroid hassasiyeti olan nefrotik sendromlu hastalarda denenmiş ve kullanılmıştır. Bu çalışmada 1uM ve 7 gün boyunca uygulanan nikotin ve levamisolün SW620 kolon kanseri hücre hattı üzerindeki etkilerini mikrodizi analizi ile gerçekleştirilmiştir. Verilerin analizi için açık kodlu R-BioConductor projesi paketlerinin yoğun olarak kullanıldığı özel bir R program tasarlanıp uygulanmıştır. Program rutini üç ana modülden oluşmaktadır: *QualCont* modülü, mikrodizilerin ve verilerin kalite kontrol analizleri için grafik destekli sonuçlar üretir ve veri normalizasyonunu gerçekleştirir; *DEGidentifier* modülü, gruplar arasındaki ifade farkı gösteren genlerin tespitini yapar ve bu konudaki sonuçları Venn şeması ve gen ifade haritası ile birlikte sunar ve son olarak

FuncAnn modülü dört ana fonksiyondan oluşmaktadır ve Gen Ontolojisi, yolak analizi bakımlarından verilerin işlevsel analiz ve açıklamalarını yapar, grafikler ve gen ağ sonuçları ile destekler.

Bu çalışma nikotin ve levamisolün SW620 hücreleri üzerinde hem gen ifade farklılaşması hem de işlevsel yolaklar bakımından oldukça benzer yanıtlar oluşturduğunu göstermektedir. Bulgularımız, nikotin ve levamisolün steroid biyosentezi, kolesterol biyosentezi ve aminoasit yıkımı metabolizmalarında oldukça güçlü bir etkiye sahip oldukları yönünde olmuştur. Mikrodizi bulguları gerçek zamanlı RT-PCR analizleri ile desteklenmiştir.

Anahtar sözcükler: nikotin, levamisol, SW620, kolon kanseri, özel R rutini, otomatik mikrodizi analizi

ACKNOWLEDGEMENTS

I would like to convey my thanks to Assoc. Prof. Özlen Konu for her supervision, guidance and support through the period since 2006 summer and throughout this study. I have learned a lot from her experience and invaluable critics. I also would like to thank her for that I learned a lot on statistics, bioinformatics and analytical evaluation of large datasets while working with. It was a privilege for me to work with her.

I would like to thank Assoc. Prof. Işık Yuluğ and Assist. Prof. Özlem İlk for their helpful comments and constructive criticisms for my thesis study.

I am especially indebted to Ceren Sucularlı, Ahmet Raşit Öztürk, Onur Kaya and Rumeysa Bıyık and Melike Öztürk who had been very good friends either in or outside the lab. I have to express that I have learned a lot from them. I should also thank to Koray Doğan Kaya for his ideas discussions.

I would like to thank Gurbet Karahan, Nilüfer Sayar, Tamer Kahraman, Chigdem Aydın Mustafa, Şerif Şentürk, Zeynep Tokçaer Keskin and Tolca Acun for their nice friendship and for helping me out with many difficulties that I encounter in the lab.

I am grateful to Vedat Taş, Hamit Taş, Alper Kılıklı and Elif Aslan for their invaluable friendship and supports, for their smiles and for sharing with me this much.

I would like to thank to Fatih Semerci and Osman Mahmut Eryurt for their invaluable friendship and supports all the years since 2003. Without them, and their friendship I would be much less than I am.

My special thanks goes to my other half Saniye, without whom I believe neither this project nor this thesis could be possible for me to finish.

Undoubtedly, my deepest gratitude would be for my whole family for their unconditional love and support for my studies and decisions.

I would like to thank TÜBİTAK for supporting me with BIDEB-2210 scholarship during my M.Sc. research period.

TABLE OF CONTENTS

LIST OF FIGURES

LIST of TABLES

1. INTRODUCTION

1.1. Nicotine

1.1.1. History of nicotine – A general overview

Although tobacco is known to have been consumed for over two thousand years, nicotine, the addictive component of tobacco, was discovered within the last two centuries. Initial isolation of nicotine dates back to 1807 yet nicotine was first purified and identified as an alkaloid with a molecular formula of $C_{10}H_{14}N_2$, by Reimann and Posselt in 1828 (Borio, 1995-2010).

In the 20^{th} century nicotine and nicotine containing compounds were used in agriculture and stock breeding as insecticides, and research records show that the scientists were interested in the effects of nicotine on nervous system, muscle irritability as well as agricultural aspects of its usage. The first study to investigate the role of nicotine in smoking habit came in 1945 (Finnegan et al., 1945). The first powerful links between smoking and lung cancer was not provided until 1950 when three different epidemiologic studies, one published in Britain and two published in the United States, demonstrated an association (Doll and Hill, 1950; Levin et al., 1950; Wynder and Graham, 1950). In 1955, Essenberg et. al. published a study on experimental lung tumor incidence in mice exposed to tobacco smoke from cigarettes low in nicotine (Essenberg et al., 1955). In the second half of the century, nicotine and cancer relationship has been one of the most investigated aspects of nicotine effect on human health. By no means less important, such popularity brought the chance of broad investigation of nicotine

along the 20th and 21st centuries in all aspects of health and disease, including toxicity, pharmacology, neurobiology, development, and even in its therapeutic usage.

1.1.2. Structure and Properties

IUPAC (International Union of Pure and Applied Chemistry) name of the nicotine is 3-(1-methyl-2-pyrrolidinyl)pyridine. It is bicyclic compound with a pyridine cycle and a pyrrolidine cycle (Figure 1)

Elemental components of nicotine are carbon, hydrogen and nitrogen, and the molecular formula of nicotine is $C_5H_4NC_4H_7NCH_3$ while the molecular weight is 162.23 grams per mole.

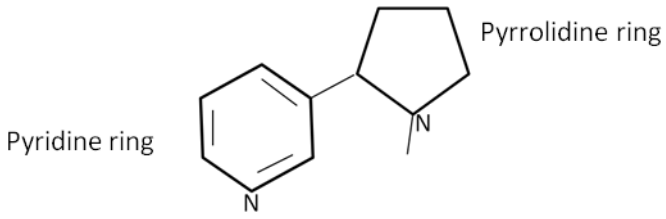

Figure 1: Chemical structure of nicotine. Adapted from (Benowitz and Jacob, 1994)

Nicotine, melting at -79°C and boiling at 247°C, is found in its liquid state at room temperature. It is miscible in water; very soluble in ethyl alcohol, ether, and chloroform (PMEP, 1985).

1.1.3. Metabolism

Nicotine is a weak base with a pKa of 8.0 (Fowler, 1954) and its absorption is reduced in its ionized form. Therefore, absorption of nicotine is dependent on the pH of the smoke: the more alkaline the smoke, the more unionized and the easier absorbed the nicotine in

mouth (Armitage et al., 1978; Brunnemann and Hoffmann, 1974; Gori et al., 1986; Sensabaugh and Cundiff, 1967). Being carried in basic tar droplets, a larger portion of the nicotine stays in unionized form which facilitates rapid pulmonary absorption (Hukkanen et al., 2005; Pankow, 2001; Pankow et al., 2003). After a puff, nicotine is rapidly absorbed in the fluid with a pH of 7.4 in the human lung and high levels of nicotine reaches to brain in 10 to 20 seconds (Benowitz, 1990; Hukkanen et al., 2005). Thus, nicotine's effects on the physiological state of an organism can be rapid and might depend on the dose and timing of the exposure.

Nicotine acts through binding nicotinic acetylcholine receptors (nAChRs) (discussed below) and exerts its effects through downstream effectors. From this route nicotine is able to enter to cell with endocytosis of desensitized receptors (Green and Millar, 1995; Kittler and Moss, 2001). On the other hand, nicotine can also enter to the cell itself; unlike for acetylcholine (ACh), membranes are readily permeable to nicotine, as it permeates at least through alveoli and capillary walls (Lester et al., 2009). Therefore, it is plausible that nicotine passes the cellular membrane also. As a matter of fact, it was shown that nicotine can promote assembly of α4β2 AChRs by inducing an active or desensitized conformation in the ER lumen (Kuryatov et al., 2005), which indicates nicotine, either permeating through the membrane or being internalized by endocytosis process, takes intracellular roles besides acting excitatory for nAChRs.

There are six major metabolites of nicotine. The most important of these is cotinine to which about 70 to 80% of the nicotine is converted (Benowitz and Jacob, 1994). In this transformation nicotine is first converted to nicotine iminium ion, by a cytochrome P450 system, primary enzyme being CYP2A6 (Murphy, 1973) and second iminium ion is

converted to cotinine by a cytoplasmic aldehyde oxidase (AOX1) (Brandange and Lindblom, 1979). Cotinine is then metabolized to *trans*-3'-hydroxycotinine again by CYP2A6 (Nakajima et al., 1996). Another primary metabolite is nicotine *N* –oxide and this conversion is catalyzed by a flavin – containing monooxygenase 3 (FMO3) (Cashman et al., 1992). A study in 1975 demonstrates that this metabolite is not further metabolized and rather is reduced back to nicotine (Dajani et al., 1975) while another indicates that this reduction takes place in large intestine (Beckett et al., 1970). These two metabolites, cotinine and nicotine *N*-oxide, are formed by oxidation of the pyrrolidine ring.

Two other metabolites are formed by a methylation and glucuronidation of the pyridine ring, respectively: nicotine isomethonium ion (*N*-methylnicotinium ion) (McKennis et al., 1963) and *N*-quaternary glucuronide (Benowitz and Jacob, 1994). Nicotine isomethonium ion formation is likely to be catalyzed by indolethylamine N-methyltransferase (INMT) enzyme, which is also known as nicotine N-methyltransferase. Dwoskin et al reported that nicotine isomethonium ion inhibits dopamine uptake in rat striatal slices and discussed that this point as to be explanatory for inverse relationship between smoking and Parkinsonism (Dwoskin et al., 1992). They further discussed that their structural relationship to neurotoxin MPP+ may be indicative of a possible protective role of N-methylated nicotine metabolites against Parkinson's Disease. *N*-quaternary glucuronide is formed via glucuronidation of the nicotine by uridine diphospate-glucuronosyltransferase (UGT) enzymes (Seaton et al., 1993) and 3 to 5% of the nicotine is known to be converted to this metabolite and excreted in urine. However, in later steps of nicotine metabolism, cotinine and *trans*-3'-

hydroxycotinine also shown to be glucuronidated and excreted in urine as cotinine *N*-glucuronide and *trans*-3'-hydroxycotinine *O*-glucuronide (Benowitz and Jacob, 1994; Byrd et al., 1992; Yamanaka et al., 2005b). Therefore, when cumulative effect is considered, glucuronidation becomes a major reaction in nicotine metabolism and enzymes catalyzing these conversions show extensive variation with differential contribution and selectivity (Berg et al., 2010; Kaivosaari et al., 2007; Kuehl and Murphy, 2003; Lessov-Schlaggar et al., 2009). UGT enzymes take part in metabolism of not only nicotine but also in elimination of other drug toxicities and xenobiotic compounds (Cecchin et al., 2009; Derby et al., 2009; Ritter, 2000; van der Bol et al., 2010) . These metabolic reactions take place in liver. Nevertheless, different tissues, particularly colon, also have expression of UGT enzymes and interestingly were found to be responsive to nicotine or drug treatments in terms of up-regulation or down-regulation of UGT enzymes (Kaivosaari et al., 2007; Kaya, 2009). Furthermore, one of these enzymes, UGT1A6, was found to be strongly up-regulated, while UGT1A1, -A3, -A4, -A5, A7, -A8, -A9 and –A10 were detectably up-regulated in SW620 colon cancer cells in response to serum starvation combined nicotine treatment (Kaya, 2009).

Fifth metabolite is nornicotine, by oxidative *N*-demethylation of nicotine (Neurath et al., 1991) and nicotine to nornicotine conversion was shown to be catalyzed by CYP2A6 and CYP2B6 in human liver (Yamanaka et al., 2005a). The last one is 2'-Hydroxynicotine and this conversion is mediated by cytochrome P450 (CYP2A6) (Hecht et al., 2000).

Although there are several carcinogenic compounds, such as nicotine and alkaloids in tobacco smoke, like N'-Nitrosonornicotine (NNN), 4-(methylnitrosamino)-1-(3-pyridyl)-

1-butanone (NNK), alcoholic derivative of the NNN (NNAL), and their metabolites (Hecht, 2002; Hoffmann et al., 2001), there is no study established to our knowledge showing carcinogenic effects of nicotine metabolites mentioned above. Furthermore, these carcinogenic compounds have been shown to be converted to non-carcinogenic metabolites and interestingly some of these derivatives are nicotine metabolites, like norcotinine and 3'-hydroxynorcotinine (Hatsukami et al., 2004; Hecht, 2002; Upadhyaya et al., 2002), which are further subjected to reactions given for nicotine metabolism. The enzymes involved in nicotine metabolism and glucuronidation are listed in Table 1 together with corresponding AffyIDs.

Table 1: Major metabolites of nicotine and enzymes that take role in catalyzing those conversions. Human Affymetrix probe identifiers are given for HGU133Plus2 Chip. Adapted from Affymetrix website.

Metabolite	**Enzyme catalyzing conversion from nicotine**	**Affy IDs for enzymes/enzyme groups on HGU133Plus2 Affy Chip**
Cotinine	CYP2A6, Aldehyde oxidase (AOX1)	**CYP2A6:** 1494_f_at, 211295_x_at, 207244_x_at, 214320_x_at **AOX1:** 205082_s_at, 205083_at
Nicotine *N*-oxide	flavin – containing monooxygenase 3 (FMO3)	206496_at, 40665_at
Nicotine isomethonium (*N*-methylnicotinium) ion	indolethylamine N-methyltransferase (INMT)	224061_at
***N*-quaternary glucuronide**	diphospate-glucuronosyltransferase (UGT) enzymes	208596_s_at, 215125_s_at, 206094_x_at, 221304_at, 221305_s_at, 207126_x_at, 204532_x_at, 236597_at, 208358_s_at, 237572_at, 232654_s_at, 206505_at, 207958_at, 211682_x_at, 235904_at, 219948_x_at, 207245_at, 207392_x_at, 228956_at, 216687_x_at, 217175_at, 232655_at
Nornicotine	CYP2A6, CYP2B6	**CYP2B6:** 217133_x_at, 206755_at, 206754_s_at
2'-Hydroxynicotine	CYP2A6	

1.1.4. NACHRS: Structure, function and localization

The nicotinic acetylcholine receptors (nAChRs) are members of a superfamily of receptors involved in ligand-gated ion transmission in both neural and muscular systems. Together with cholinergic muscarinic receptors, neuronal nAChRs have central role in mediating ACh message transduction. Their role in mediating fast actions of ACh in neuromuscular junctions and nervous system is well defined (Arias, 1997; Changeux and Edelstein, 1998; Gaimarri et al., 2007; Gotti and Clementi, 2004; Lena and Changeux, 1998). They have widespread expression in nervous system and they transduce cholinergic signals in many brain areas and peripheral ganglia. nAChRs mainly have pre-synaptic localization in central nervous system and modulate neurotransmitter release, nevertheless in some areas they are localized post-synaptically and mediate fast synaptic transmission (Dajas-Bailador and Wonnacott, 2004; Gotti and Clementi, 2004; Jensen et al., 2005). Although originally classified as "neuronal" and "muscular", later in time they are found to be expressed in a wide range of tissues including muscle, lymphoid tissue, macrophages, skin, lung cells, vascular tissue, and astrocytes (Gotti and Clementi, 2004). Additionally, in previous studies we have shown colonic expression, in human SW620 colorectal cancer cells, and liver expression in zebrafish (unpublished data).

nAChRs are composed of homo- or heteropentamers of subunits, which are classified to five groups: α, β, γ, δ and ε. γ, δ and ε subunits are usually thought to be auxiliary subunits and a nAChR is denominated according to its α and β content on the grounds that these two subunits are important for pharmacological specificity and sensitivity of a complete receptor (Luetje and Patrick, 1991). Ten α-type (α1-α10) and four β-type (β1-

β4) subunits are identified in human to date, among which α1 and β1 are muscle type subunits and others are classified to be of neuronal type. All subunits have four transmembrane domains with an intracellular loop between third and fourth, and both amino- and carboxy-terminal portions are localized extracellularly.

Five subunits are needed to build up a fully functional receptor, and receptor composition and subunit stochiometry is decisive for specificity and functionality. Only α7 and α9 subunits are capable of forming homopentamers, on the other hand receptors containing α2-α6 and β2-β4 subunits form only heteromeric receptors (Gotti and Clementi, 2004; Lindstrom, 2000).

1.1.5. nAChRs and Cancer in Human

Cholinergic signaling and nAChR involvement in several cancer types have been shown in previous studies (Schuller, 2009), since these receptors regulate neurotransmitter release which in turn might activate release of different growth factors. For instance, α7 homomeric receptors, shown not to be desensitized unlike heteromeric receptors (Kawai and Berg, 2001), release adrenaline and noradrenaline which in turn increases epidermal growth factor (Carlisle et al.), vascular endothelial growth factor (VEGF) and arachidonic acid (AA) (Heeschen et al., 2002; Wong et al., 2007). On the other hand, α4β2 heteromeric receptors stimulate γ-aminobutyric acid (GABA) release (Al-Wadei and Schuller, 2009), which has been shown to inhibit tumor growth in colon (Joseph et al., 2002), breast (Drell et al., 2003), lung (Schuller et al., 2008b) and pancreas (Schuller et al., 2008a).

There is not much known about the association of muscle type nAChRs, α1 and β1, with cancer. Only one study has shown increased expression of these subunits in non-small cell lung cancer (NSCLC) (Carlisle et al., 2007). However, both of these subunits are involved in congenital myasthenic syndrome with mutations (Garchon et al., 1994; Quiram et al., 1999).

1.1.6. Nicotine and SW620 Cells

Previously, our group have investigated the effects of nicotine on colon cancer cells under serum deprivation (0.1% FBS) and normal growth conditions (Kaya, 2009). In that study it was shown that, when treated with serum deprivation, 1μM nicotine affected KEGG pathways such as cell cycle, calcium signaling and MAPK, cell adhesion and cell communication pathways, glucuronidation metabolism, coagulation cascade, ribosomal genes and purine-pyrimidine metabolism genes. Furthermore, that study have revealed that nicotine had the ability to relieve the cell proliferation from serum deprivation-induced suppression and saves the cells from starvation-induced apoptosis. However, that study also have shown that nicotine when given to cells under normal growth conditions (10% FBS) for 2 days did not demonstrate such an effect. A 2 day treatment can perhaps be considered acute whereas 7 and 14 day treatments are rather chronic.

1.2. Levamisole

1.2.1. Structure, Properties and Antihelminthic Mode of Action

Levamisole, L-isomer of tetramisole, was originally discovered at Janssen Pharmaceutica in 1966 and it is a synthetic derivative of imidazothiazole

(HelminthInfections, 2010). Figure 2 is a drawing representing the chemical structure of levamisole. Levamisole is also a weak base, like nicotine, with pKa of 8.0 and generally sold as hydrochloride salt form on the market with a molecular weight of 241 grams per mole (Yakoub et al., 1995). Additionally, it is light sensitive and stable under ordinary conditions. It melts at about 230 to 233^{o}C and is soluble in water (Yakoub et al., 1995).

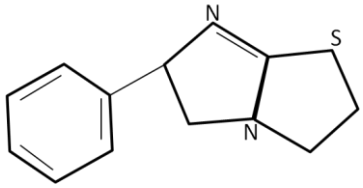

Figure 2: Chemical structure of levamisole. Adapted from (Roberts, IPCS INCHEM)

Levamisole has a broad use as an antihelminthic in livestocks since it eliminates parasitic worms by paralysis. Other than the agricultural use, given the fact that over 1 billion people in the developing world are infected with intestinal nematodes (Bethony et al., 2006), levamisole is also important in solving public health problems.

Levamisole binds to acetylcholine receptors and causes body-wall muscle hypercontraction, spastic paralysis, and ultimately death of parasites (Lewis et al., 1987a; Lewis et al., 1987b). Levamisole is reported to be sensitive to receptors on nematode and that it does not activate the AChRs on the host (Rayes et al., 2004), and such difference was thought to be due, nevertheless biophysical similarities of receptors, to the pharmacological differences between the nematode and mammalian host. However, molecular basis of such specificity was not elucidated enough. In a more recent study, Boulin et al showed that levamisole-specific AChRs were different from those bind nicotine (Boulin et al., 2008) and this may explain the specificity as to be not at the species level but at the molecular level. Nicotinic and levamisole-specific acetylcholine receptors will be discussed in detail later.

1.2.2. Metabolism

Levamisole was found to have a fast absorption and metabolism in humans after oral administration so that unchanged drug peak was reported to be reached 2h after dosing of ^{3}H-levamisole while at this time about 60% of the radioactivity was reported to be found in the form of metabolites (Roberts, IPCS INCHEM). Levamisole is extensively metabolized in humans and only about 4 to 5% is excreted unchanged. Experiments on levamisole metabolism were performed generally on isolated hepatocytes of dogs, sheep, cattle, pigs and human; and some on rats and monkeys. One of the two major pathways is dehydrogenation and subsequent sulfoxidation of the imidazolidine ring, while the second being aromatic hydroxylation. This p-hydroxy levamisole has a glucuronide conjugate and together they make up to 17% of the administered levamisole. These experiments were performed in Janssen Research Foundation and submitted to World Health Organization (WHO) as unpublished experimental and preclinical reports (Roberts, IPCS INCHEM).

1.2.3. L-AChRs: Structure, function, localization

L-AChRs (levamisole-sensitive AChRs) are acetylcholine receptors sensitive to levamisole and similar to nAChRs, they are ligand gated ion channels functioning as pentamers. These receptors have greatly been studied in worm species, especially in nematode *C. elegans* where most of the information about them came from. *C. elegans* has the most extensive subunit variation of acetylcholine receptors with 27 subunits identified and provided on WormBase, 10 of which were identified from levamisole-resistance or neuronal degeneration studies and 17 from functional genomics studies

after completion of its genome sequencing: *unc-63/lev-7, unc-38, deg-3, des-2/acr4, acr5-12* and *acr15-24* are alpha type; *unc-29, lev-1, acr2, acr3, acr14* and *acr25* are non-alpha type subunits (Ballivet et al., 1996; Baylis et al., 1997; Fleming et al., 1997; Jones and Sattelle, 2004; Mongan et al., 1998; Mongan et al., 2002; Squire et al., 1995; Treinin and Chalfie, 1995; Treinin et al., 1998; WormBase, 2010). Other than these AChR subunits, *ric-3* (Halevi et al., 2002; Halevi et al., 2003), *unc-50* (Eimer et al., 2007) and *unc-74* (Lewis et al., 1980) are found to be required for in vivo expression of L-AChRs.

In *C. elegans* neuromuscular junctions, excitatory neurotransmission is mediated by L-AChRs and UNC-29, UNC-38 and UNC-63 were found to be absolutely required for a functionally active receptor (Boulin et al., 2008; Culetto et al., 2004; Fleming et al., 1997; Lewis et al., 1987b; Rayes et al., 2007).

One of the key points for safe usage of levamisole as medical and agricultural antihelmintic is the discrimination between the parasite and the host receptors. Acetylcholine receptors of human and other mammalians were investigated for levamisole sensitivity and levamisole was found to be a weak agonist of human nAChRs (Levandoski et al., 2003; Rayes et al., 2004). Another study revealed that nicotine-sensitive and levamisole-sensitive receptors of *C. elegans* were insensitive to the ligands of each other (Boulin et al., 2008). Taken together, the data raised questions about the efficacy of levamisole in medical considerations other than soil transmitted helminth infections and antiparasitic in livestocks.

Two most important of medical conditions treatment with levamisole might be relevant to are nephrotic syndrome and adjuvant treatment of colon cancer in combination with 5' fluorouracil. While levamisole could not strongly activate human nAChRs, how it could have systemic effects, such as remission of edema, oliguria (Dayal et al., 1994), disappearance of proteinuria in nephrotic syndrome (Rashid et al., 1996), increase in disease free and overall survival in colorectal cancer (Laurie et al., 1989; Moertel et al., 1990), in human is still a mystery.

Furthermore, despite levamisole effect on these diseases in human was successfully shown in many studies, limitations of both success-showing and failure-showing studies and additional inconsistencies between the limitations of those studies has made the issue much more complicated. Therefore, research of how levamisole exerts its effect on mammalian cells and systems has been overshadowed among the debate of presence or absence of such effect.

1.2.4. Levamisole and Apoptosis

In addition to antihelmintic activity of levamisole, it was also shown to act as an apoptosis inducer. Investigation of such a role for levamisole was motivated by its usage in colorectal cancer chemotherapy, which will be discussed later in this study, as an adjuvant in combination with 5-Fluorouracil. One of the most detailed studies came from Artwohl et al in 2000 (Artwohl et al., 2000). In this study, levamisole at 0.5 mM, 1 mM, 2 mM and 5 mM concentration levels was shown to be a strong apoptotic inducer in human umbilical vein endothelial cells, adult human venous endothelial cells, and human uterine microvascular endothelial cells; but not in fibroblasts. One study with

human hepatoblastoma (HepG2) cells shows that levamisole, by being an activator of cystic fibrosis transmembrane conductance regulator (CFTR) Cl^- channels, inhibits the apoptosis induction by glibenclamide (Kim et al., 2001), which was also shown to induce apoptosis by inhibiting CTFR Cl^- channels in HepG2 cells (Kim et al., 1999). Taken together, levamisole might have a controversial association with apoptosis and/or apoptotic machinery. It appears levamisole is actually not an apoptotic or anti-apoptotic agent *per se*, but may be involved in the process in a very indirect manner depending on the context of apoptosis (intrinsic or extrinsic pathways), on the receptor status of the cell for levamisole, and on the dose of application, and finally on the cell type itself. Such an interpretation is also in line with the fact that there is no study established to our knowledge that could able to show a more direct interaction between any of the components of central apoptotic machineries and levamisole.

1.2.5. Levamisole, Colorectal Cancer and Cell Proliferation

Levamisole was considered for adjuvant chemotherapy for colon cancer in combination with 5'Fluorouracil and it this chemotherapy was offered as standard for especially late stage colorectal cancer cases by National Cancer Institute Consensus Conference in 1990 (NIH_Consensus_Conference, 1990). Studies for levamisole usage in colorectal cancer therapy started after 1990 and have mostly been based on two studies, preceding the NIH Consensus Conference (Laurie et al., 1989; Moertel et al., 1990). Levamisole has been tried for immunomodulatory effect in colorectal cancer since its immune enhancing effect had been appreciated already (Dahl et al., 2009; Laurie et al., 1989; Leibovici et al., 2009). Levamisole's effect on cell proliferation is also based on findings showing an enhancement of Th1 immune response and increased number of natural killer (NK) cells

(Holcombe et al., 1998; Zhang et al., 2009). However, there had been many studies unable to verify an efficient role for levamisole either in disease free survival (DFS) or overall survival (OAS) in colorectal cancer (Cascinu et al., 2003; De Placido et al., 2005; O'Connell et al., 1998). At the end, the role of levamisole in colon cancer treatment remained inconclusive, and it was withdrawn from clinical usage. After 2003, Jansen-Cilag, pharmaceutical company that discovered levamisole in 1960s and produced since then, decided to stop production of ergamisol, the brand name for levamisole, for clinical usage (Davin and Merkus, 2005). Indeed, there is a need to decipher the role of levamisole, which had been frequently used as an adjuvant therapy, without knowing downstream molecular effects.

1.2.6. Levamisole and Nephrotic Syndrome

Levamisole has been tried for steroid-sparing effect in nephrotic syndrome (Davin and Merkus, 2005; Tanphaichitr et al., 1980), which is a disorder with protein loss from blood to urine, with damaged kidneys. Since nephrotic syndrome treatment widely includes corticosteroids, with adverse effects in steroid-sensitive patients with nephrotic syndrome. Based on the study by Tanphaichitr et. al in 1980, levamisole was studied for sparing steroid from disease treatment, in other words, for reducing the steroid content of treatment. Although this issue has also remained inconclusive and overshadowed by the debates of presence or absence of a sufficient effect, there has not been any study to date questioning why levamisole should have such effect if any, and what could be the relationship between steroid metabolism and levamisole.

1.3. Microarray Analysis Methods

1.3.1. Microarray Analysis Programs Based on R Environment

Microarray studies have become widespread and enormous amounts of data have emerged and continue to do so. It is crucial for a transciptomics dataset to be correctly preprocessed, analyzed and interpreted since the quality of these steps has decisive role in making the findings of the study meaningful. The flourishing of microarray experiments as a high throughput method has brought the necessity of designing tools for helping in handling and statistical evaluation of such large datasets, and functional interpretation of biologically relevant findings.

R, an open-source programming language and environment developed for statistical computing and graphics, provides a nice opportunity and beneficial means for analysis of microarray data (R_Development_Core_Team, 2010). Nevertheless, R has a script based environment and its interface, like all other programming environments, might be difficult to handle for inexperienced users without adequate mathematical, statistical and computational background. Such difficulties create the need for design and production of customized routines and/or software to ease analysis steps. Therefore, R has a unique place since it contains a large array of packages geared for microarray preprocessing and analysis, although others exist that are based on C, Java, and Matlab®.

Most of the web based services designed to perform multistep microarray analysis service provide mostly the statistical analysis of the data, with fixed parameters or with limited parametric options (http://crcview.hegroup.org, http://sourceforge.net/projects/geckoe). The quality control assessments of the arrays may

always not be included in such analyses; yet quality control is of special importance because if any of the arrays is poor in quality then it exerts a bias on the following statistical analyses. Such shortcomings and fragmentation of the whole analysis makes working on such high throughput datasets tedious; thus customized routines written in R and geared toward the user's own experimental design and analysis requirements are usually needed.

One of the array analysis software based on R is BRB Array Tools, developed by Richard Simon and BRB-Array Tools Development Team, and works as an add-in with Microsoft© Excel® interface. This tool utilizes an Excel® frontend, which cannot be run under a Linux/Unix system, whose use might provide additional flexibility in integration of customized R routines into the analysis packages and larger memory options. In addition, functional analyses provided by BRB tools do not include gene network visualization. Other efficient R-based tools include Bionforx (http://bioinforx.com/), which is proprietary, WebArray (Wang et al., 2009; Xia et al., 2005; Xia et al., 2009), EMAAS (https://www.emaas.org/EMAAS/), CARMAweb (Rainer et al., 2006); these serve the users online. Finally Babelomics team also employ a comprehensive microarray analysis service in their website based on analysis of gene lists, with Babelomics version 4.0 (www.babelomics.org). All of these tools and services are based on R environment and make use of BioConductor Project packages, especially in normalization and statistical analysis steps. Although such tools available and highly useful, having a customized R-routine to the user's experimental design, which is also modular and expandable, makes possible to add the most appropriate statistical design,

visualization and analysis tools as well as the recent algorithms and packages available in R environment.

1.3.2. BioConductor

BioConductor (www.bioconductor.org) is a software project with open source and open development features, and designed to produce tools for high-throughput genomic data analysis (Gentleman et al., 2004). Being primarily based on R programming language, most important part of the project is that it is composed of purposive R-packages, which can be downloaded via BioConductor's website and installed into R library. By definition a 'package' is designed to include related software components and documentation for purposive objectives. Packaging system is well established in R and provides independent yet interoperable modules for usage in high-throughput data analysis to achieve various statistical analytic tasks and for visualization purposes. On the other hand, BioConductor provides annotation data packages, which ease mappings between different identifiers, such as affy IDs and Entrez genes, and such service enables association of array data to biological metadata from web databases. These annotation tools are beneficial for assembling and processing genomic annotation data from GenBank, Entrez genes, Unigene, Gene Ontology (GO) Consortium and USCS Human Genome Project. Since Bioconductor is an open source environment, microarray analysis packages available within the environment that range from quality control to differential expression and to functional analyses, can be easily integrated into the custom routines written in R.

1.3.3. Functional Analysis of Microarrays Using Online Software

Once microarray data has undergone processing procedures and statistical analyses, functional analyses of the genes that have been identified to be differentially expressed should be performed. This step is of crucial importance on the grounds that only then the holistic impact of the experiment on the cellular response could be evaluated. Questions regarding the cellular processes and pathways that are influenced from the experiment find their answers with the comprehensive functional analyses that allow a more complete understanding of the nature of the experiment.

There are several online software tools concerning functional analysis, all of which incorporate either Gene Ontology (GO) information or Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways or enrichment analysis with predefined or customized gene sets. Some of these services are provided by DAVID (http://david.abcc.ncifcrf.gov/), WebGestalt (http://bioinfo.vanderbilt.edu/webgestalt/), PathwayMiner (http://www.biorag.org/), GSEA (http://www.broadinstitute.org/gsea/), GSEA Molecular Signature Database (http://www.broadinstitute.org/gsea/msigdb/), Babelomics (www.babelomics.org), Ingenuity (http://www.ingenuity.com/).

2. AIM of THE STUDY

This study uses microarray analyses methods to understand the differences and similarities between the levamisole and nicotine, two of the cholinergic agents with medical implications, on the SW620 colon cancer cells.

Here in this study we aim to assess the influences of these two agonists of nAChRs on human cells in terms of transcriptome profiling and the interpretation of functional implications and indications based on their differential and similar responses. Although thought as a therapeutic agent in several human disorders including colorectal cancer and nephrotic syndrome, levamisole has previously not been investigated for its effect at the cellular level. Furthermore, there exists no study established to our knowledge assessing the effects of levamisole on global transcriptional profile of the cells. Given the fact that efficiency studies on disorders, for which levamisole has been concerned as a therapeutic agent, remained inconclusive, we believe that our study will shed much light to comprehend an exact role for levamisole in treatment of those disorders at the dose specified in our experiments. On the other hand, we approached to the issue with a comparative manner and we compared and contrasted the effects of levamisole with nicotine; this approach is sound because inasmuch as both compounds are agonists of nAChRs, there is not a levamisole-sensitive receptor subset in humans. Furthermore, nicotine is known to have different effects upon acute or chronic exposure. Previous studies (Onur Kaya, 2009 Ms. Thesis) in our lab showed that a short term treatment of SW620 cells with 1μM nicotine did not result in large numbers of genes being differentially regulated. Therefore, we set out to test the effects of nicotine and

levamisole for an extended period of 7 days at which the exposure could be called chronic (Pakkanen et al., 2006; Sun et al., 2003).

For data analysis, a customized R-routine was written using the available resources in BioConductor and related R packages to perform an automated analysis of one-factor experimental design of Affymetrix arrays (or any single-channel data) with two or more groups; the aim was to perform the analyses from the very beginning to the end, seamlessly; so the steps included routines for array quality control assessment, differential expression, functional analysis and annotation, in a streamlined manner. We provide a generic implementation to enable the usage of the routine in any expression data independent of the number of groups. Results are generated in R and saved to a user defined directory with a given denomination extracted from the experiment descriptor file. Accordingly, we aim to provide a useful, strong and easily handled and expandable tool for researchers inexperienced in the analysis of large datasets; for the experienced ones we aim to provide additional flexibility with several options in analysis parameters.

In general, our study aims to *i)* provide a streamlined routine for the analysis of one-factor experimental design of Affymetrix arrays using R programming language; *ii)* to generate list of genes similarly modulated by nicotine and levamisole; *iii)* to generate list of genes differentially modulated by nicotine and levamisole; *iv)* to perform functional analyses on these lists by the provided routine and finally *v)* to confirm the expression pattern obtained from array data of selected genes using real-time polymerase chain reaction (PCR) in SW620 cells exposed to levamisole or nicotine.

3. MATERIALS AND METHODS

3.1. Levamisol and Nicotine Treatment of Cells for Microarray and RT-PCR analysis

SW620 colon cancer cells were seeded into the 75 cm^2 flasks with a density of $1x10^6$ cells per flask in DMEM supplemented with 10% fetal bovine serum (FBS) and 0.1% penicillin/streptomycin (P/S) mixture. Cells left for 24h for attachment to the plate and then medium was replaced with 0.1% FBS-containing medium to make cells quiescent (synchronization). Media then were replaced with fresh media that contain 1 μM nicotine- or 1 μM levamisole Media with treatments were replenished every 3 days. Cells incubated at 37^oC with 5% CO_2. Cells were gathered from the plates at day 7 or 15. For detaching cells from plate, trypsin/EDTA was used which was then neutralized by growth medium. Cells were centrifuged at 1200 rpm for 3 minutes, washed with PBS and stored at -80^oC. Microarray optimizations, cell culture experiments and sample preparations were performed by Onur Kaya, a former MS student at Bilkent University.

3.2. RNA extraction, cDNA synthesis and real-time PCR experiments

RNA extraction was performed with Promega SV Total RNA Isolation Kit according to the manufacturers protocols except for some of the cases, a second DNAse treatment step preceding elution was added (Z3100; Madison, USA). cDNA synthesis was performed with Fermentas RevertAid™ First Strand cDNA Synthesis Kit (K1631; Lithuania). For confirmatory real-time PCR experiments, DyNAmo HS SYBR Green qPCR Kit from Finnzymes (F410L, Espoo, Finland) was used and experiments were

carried on Bio-Rad iCycler machine. Primers were designed with Primer 3.0 (Rozen and Skaletsky, 2000). Real time PCR conditions for all primers were as follows: 95°C 10 min; 45 cycles 95°C 30 s, 60°C 30s, 72°C 30 s; 72°C 10 min. Real-time PCR experiment data were analyzed with Pfaffl method: $ratio=[(E_{target})^{deltaCt\ target\ (control-treated)} / (E_{ref})^{deltaCt\ ref\ (control-treated)}]$ (Pfaffl, 2001); reference gene was chosen as beta-actin, which showed no difference both in our microarray data and real time PCR experiments.

Table 2: Sequences and efficiencies of the primers of the genes selected for confirmatory real-time PCR experiments.

Gene	Primer Sequence	Primer Efficiency
ACTB	5' CCAACCGCGAGACGATGACC 3'	1.99
	5' GGAGTCCATCACGATGCCAG 3'	
GULP1	5' CAGGCAGTATGACACCTAAG 3'	2.34
	5' CAGGTCCCGTTTAATCTCAG 3'	
CHRNA1	5' GGCATCAAGTACATCGCAGA 3'	2.18
	5' TCAATGAGTCGACCTGCAAA 3'	
CHRNB1	5' CCTGACGTGGTGCTACTGAA 3'	2.18
	5' CAGCTGCTGCGATAGATGC 3'	
MKI67	5' GTGTCAAGAGGTGTGCAGAA 3'	2.28
	5' GCCTTACTTACAGAATTCAC 3'	
XIAP	5' TCACTTGAGGTTCTGGTTGC 3'	2.22
	5' CGCCTTAGCTGCTCTTCAGT 3'	
FOSL2	5' GGCCCAGTGTGCAAGATTAGCC 3'	2.03
	5' TTTCACCACTACAGCGCCCACC 3'	
MAP1B	5' GTTGGAAGGAAAGGCTCAGT 3'	1.99
	5' CTTGCTGTTTCTCATGGGTC 3'	

3.3. Microarray experiments

Affymetrix GeneChip® U133 Plus2 arrays were used for sample hybridization; and 5µg RNA were used for each experiment (amplification/labeling/hybridization). Three different conditions were used for microarray experiments: 1) 1µM 7 days nicotine treatment, 2) 1µM 7 days levamisole treatment and 3) control; each group included two

replicates. Amplification, labeling and hybridizations were performed at the Genomics Core Facility of Bilkent university by the facility technician Bilge Kılıç under the supervision of Assoc. Prof. Dr. Işık Yuluğ according to the manufacturer's protocols. GeneChip Operating Software was used for preliminary probe-level quantification as previously described (Kaya, 2009).

3.4. Quality Controls and Pre-processing of the microarray data

For quality control and normalization of the data and for the determination of the differentially expressed genes followed by functional analysis, a streamlined R-routine was written (See Appendix). This R-routine basically has been composed of three different modules, first of which performs the quality control steps and data normalization (*QualCont*) while the second identifies differentially expressed genes (*DEGidentifier*) and the third performs functional annotation and analysis of the results (*FuncAnn*) (Figure 3). All three modules extensively use BioConductor packages (Figure 4) (www.bioconductor.org).

QualCont reads the data in from .CEL files provided with a phenodata file, an experiment descriptor file created by the user in a tab delimited text format, uploaded from a given directory. *QualCont* applies to the raw dataset a user-selected normalization method (i.e., RMA, gcRMA or MAS5) using the `affy` package (Gautier et al., 2004) and generates files containing normalized results in a spreadsheet as well as image formats. The function also assigns some of the results into the global environment of R to make some of the results readily available to the user for further analyses in the following modules. The R packages used in the *QualCont* function are `affy` (Gautier et

al., 2004) and `affyPLM` (Bolstad, 2004; Bolstad et al., 2005; Brettschneider et al., 2007). *QualCont* function outputs include: photo images of the chips, normalized data files, histogram representation of the data distribution for the raw and normalized data, MA plots of the raw and normalized data and boxplots of raw and normalized data, and RNA degradation plot of the data (Figure 4). The function saves all these information in the directory provided by the user as the primary input to be used in other functions.

3.5. Determination of Differentially Expressed Genes

Differentially expressed genes were identified by the second function of the R-routine, *DEGidentifier*. This function has been designed as to apply a linear model and Bayesian statistics for the assessment of differential expression for one-factor randomized design experiment using the limma package in R (Smyth, 2005).

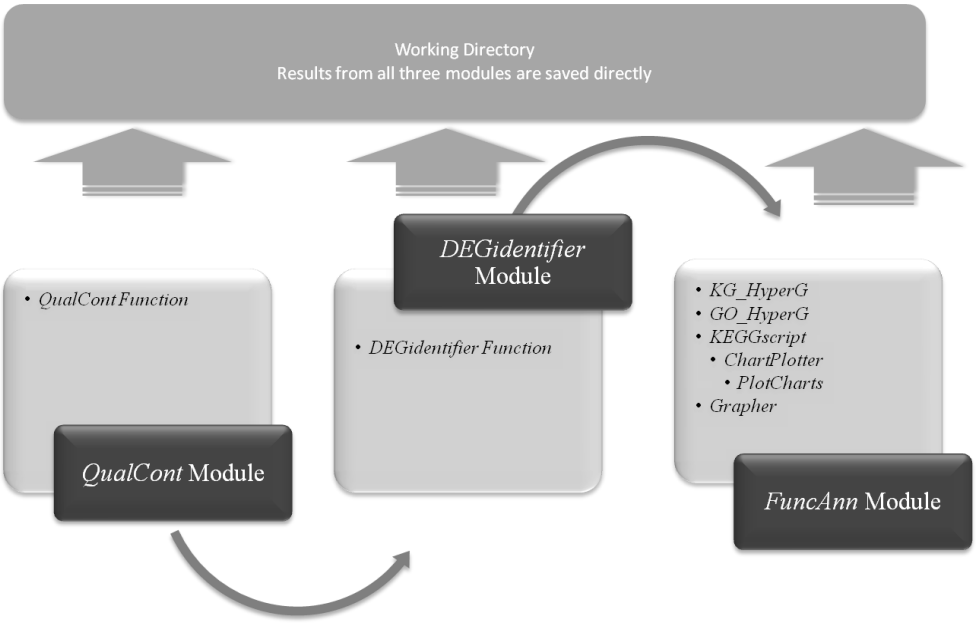

Figure 3: Workflow of the routine. The routine is composed of three modules, each of which is composed of related functions. At each step, results are automatically saved to the user-defined directory, with appropriate denomination.

Accordingly, the contrast groups are formed also by the function based on the experiment description phenodata file, location of which is provided by the user. Normalized data are then fitted to a linear model for each gene, and estimated coefficients and standard errors are computed and finally differentially expressed genes are determined by empirical Bayesian methods. The R packages used in *DEGidentifier* module are: `affy` (Gautier et al., 2004), `limma` (Smyth, 2005) and `hgu133plus2.db` (Carlson et al., R package version 2.4.1).

DEGidentifier provides differentially expressed genes between all paired groups in a given dataset and saves the retrieved information in different spreadsheets to a user defined directory. Outputs of the function are as follows: All probesets with significant differential expression between any two groups, separate up- or down-regulation lists for probesets, log-fold change of the intensities between groups. Additionally, it generates all the lists given above at the gene symbol level as well as at probe level, and creates Venn diagram and Heatmap representations for the comparisons between groups (Figure 13, Figure 14).

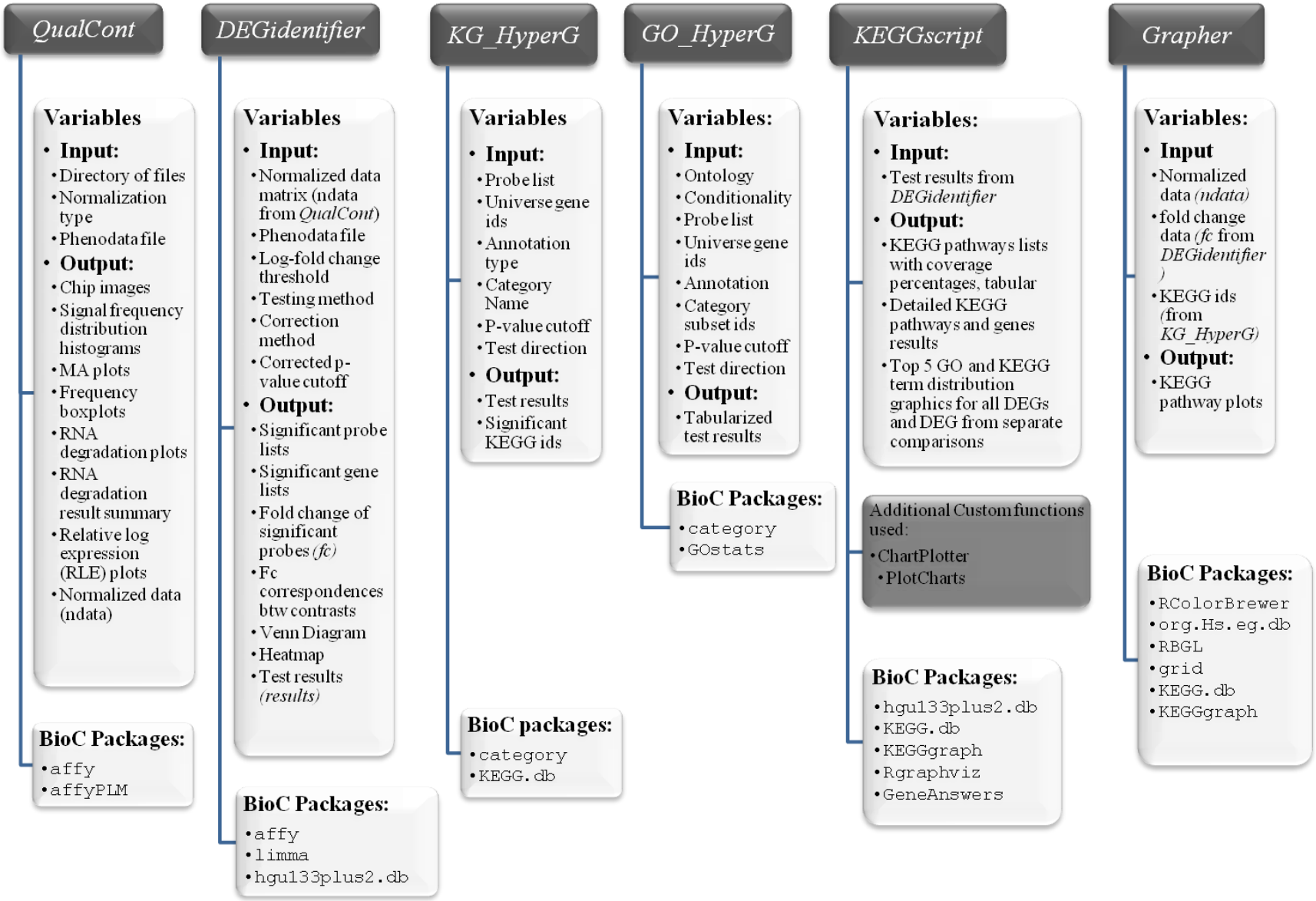

Figure 4: Detailed schematic view of routine components. Input values and output results for each function as well as used BioConductor packages are given. Additional Custom functions *ChartPlotter* and *PlotCharts* are two functions modified from equivalent functions of package `GeneAnswers` (Feng et al., 2001). Those modifications were necessary for us to ensure streamline flow of the routine.

3.6. Functional Annotation and Analysis of Differentially Expressed Genes

The third module of the R-routine, *FuncAnn*, performs functional annotation for differentially expressed genes in terms of Gene Ontology and KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathways (Figure 17-Figure 21 and Figure 24; Table 6-Table 9). *FuncAnn* is composed of four different functions, *KG_HyperG, GO_HyperG, KEGGscript* and *Grapher*, written by us and two functions, *ChartPlotter* and *PlotCharts*, modified from the `GeneAnswers` package of BioConductor (Feng et al., 2001). This module makes use of following BioConductor packages: `KEGG.db` (Carlson et al.), `KEGGgraph` (Zhang and Wiemann, 2009), `GeneAnswers` (Feng et al., 2001) and `RGraphviz` (Gentry et al.).

KG_HyperG provides the results of hypergeometric testing of KEGG pathways in which the DEGs are involved. *GO_HyperG* function performs hypergeometric testing of GO terms for overrepresentation or underrepresentation of GO terms in a given significant probe list, independent of the chip platform as long as it is an Affymetrix human chip, in a given universal set of probe ids (default is all the probes defined on the chip). *KEGGscript* function generates outputs of detailed KEGG analyses of the results, separately for each contrast, where the name of the pathway, the coverage percentage of the pathway, the names of genes differentially expressed in that pathway and whether those genes are up-regulated or down-regulated are provided as output, by making use of `KEGGgraph` package of BioConductor. A table with the same information also is generated and saved to the working directory in spreadsheet format. Furthermore, we included graphical results for gene annotation within *KEGGscript* function. Top GO

processes and top KEGG pathways, to which the data are related, are provided as piechart and barplot, with the help of `GeneAnswers` package. Finally, *Grapher* function generates graphical results for KEGG pathways with directed edges and color code for up-regulation and down-regulation using `KEGGgraph` package of BioConductor.

All the output of this module also is automatically saved into the working directory in easily readable plain text formats, spreadsheet formats or as images; all automatically denominated in accordance with the contrast groups (e.g.,. "levamisole-nicotine_KEGGPathways.txt").

4. RESULTS

4.1. Quality Control Assessment of Hybridization and Preprocessing of Microarray Data

The *QualCont* program was applied to 6 microarray .CEL files that belong to 3 groups, nicotine, levamisole, and control. Each group contained 2 samples.

```
QualCont (FilesPath,NormType,pdfile)
```

Here FilesPath is the directory where .CEL files and experiment descriptor file (pdfile) are located. Since the function was designed as to read all available .CEL files in the given directory, it is important to create a separate folder for separate analyses. Furthermore, the directory given here is set as the working directory of the session, all the results are saved here from the beginning till the very end of the analysis. NormType is the normalization type (RMA, gcRMA or MAS5; all with lower letters) selected; and herein "rma" (Robust Multi-array Analysis) (Irizarry et al., 2003) was chosen as the normalization method. Pdfile is a descriptive text file explaining the experimental groups.

Quality control analyses showed that there was no physical defect the on hybridized chips, like smears, (Figure 5) and all samples had similar RNA degradation slopes (Figure 10), indicating that all the arrays were of adequate quality to be included in future analyses (Table 3).

Histograms (Figure 6, Figure 7), boxplots (Figure 8) and MA plots (Figure 9) of raw and normalized data were provided for comparison. Histograms demonstrated the signal

intensity distribution for the arrays. Post-normalization the distributions were more similar to normal distribution yet with a heavier right tail (Figure 7). Here we provide histograms, however Q-Q plots can also be very informative to see whether the signal intensity has normal distribution or not (R_Development_Core_Team, 2010). We should note here that for *t-test* assumptions, meeting of normality in distribution of signal intensities of "most probesets" across experiments is more important than satisfying normality of distribution of a single array, since we apply testing "between" arrays rather than "in" an array. Nevertheless, we found it useful to provide in-array-distributions of signal intensities to get an idea about how general overview of signal distributions are seen in arrays. Additionally, it would be beneficial to keep in mind that with very low numbers of arrays (e.g. 6 in our data) it would not be easily possible to prove normality, since data with low sample size generally deviate from normal distribution. Thus use of tests such as limma are better suited for small sample size comparisons. Histogram representations were generated, denominated and saved automatically by the routine and `hist` function was used in the code.

Another quality control measure is provided with the MA (MvA; M versus A) plots where one can get a quick overview of the intensity-dependent ratios of arrays.

The basic assumption of microarray theory is that most of the genes do not show differential expression between the groups. For testing whether this assumption applies to a data we look for the ratios of signal intensities in arrays as pairs. According to the aforementioned rule, we expect to have mean of intensity ratios between arrays as *1* (*0*, when logged; *log(1)=0*). In a paired MA plot, this rule is visually checked as a quality control on the grounds that if there is much difference between two arrays, in other

words if the intensity ratio between arrays do not approximate to 1, that means the most of the genes in one of them shows differential expression. Such a result implies that the given array should be excluded from the analysis because it exerts a bias on the normalization procedure, which in turn leads to a bias in the statistical analytic steps.

Therefore, in the streamline routine we included generation of MA plots of both raw and normalized data and we made use of `mva.pairs` function of `affyPLM` package of BioConductor.

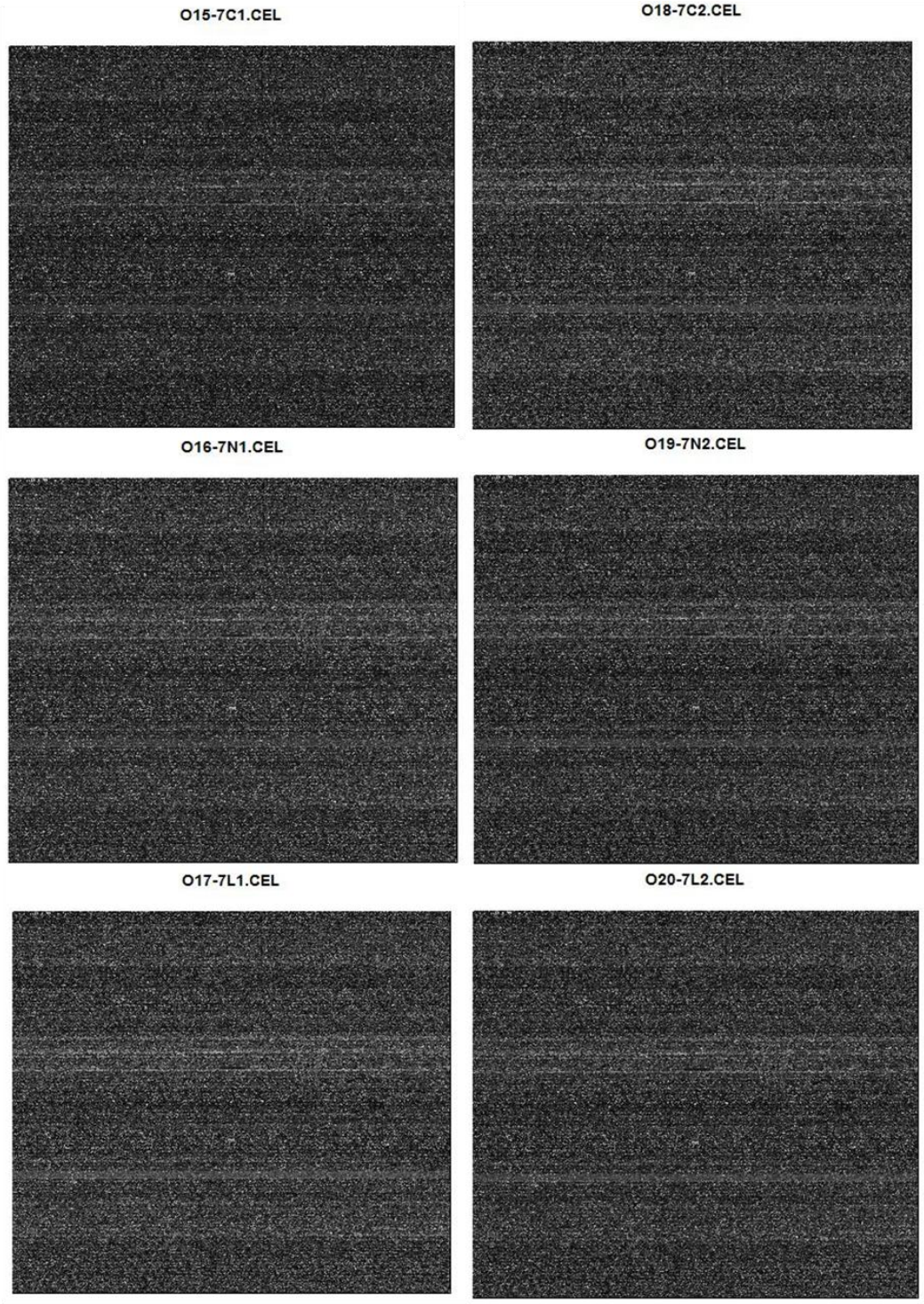

Figure 5: Photographs of array chips. Images were created with the help of `image` function of the `graphics` package (R_Development_Core_Team, 2010).

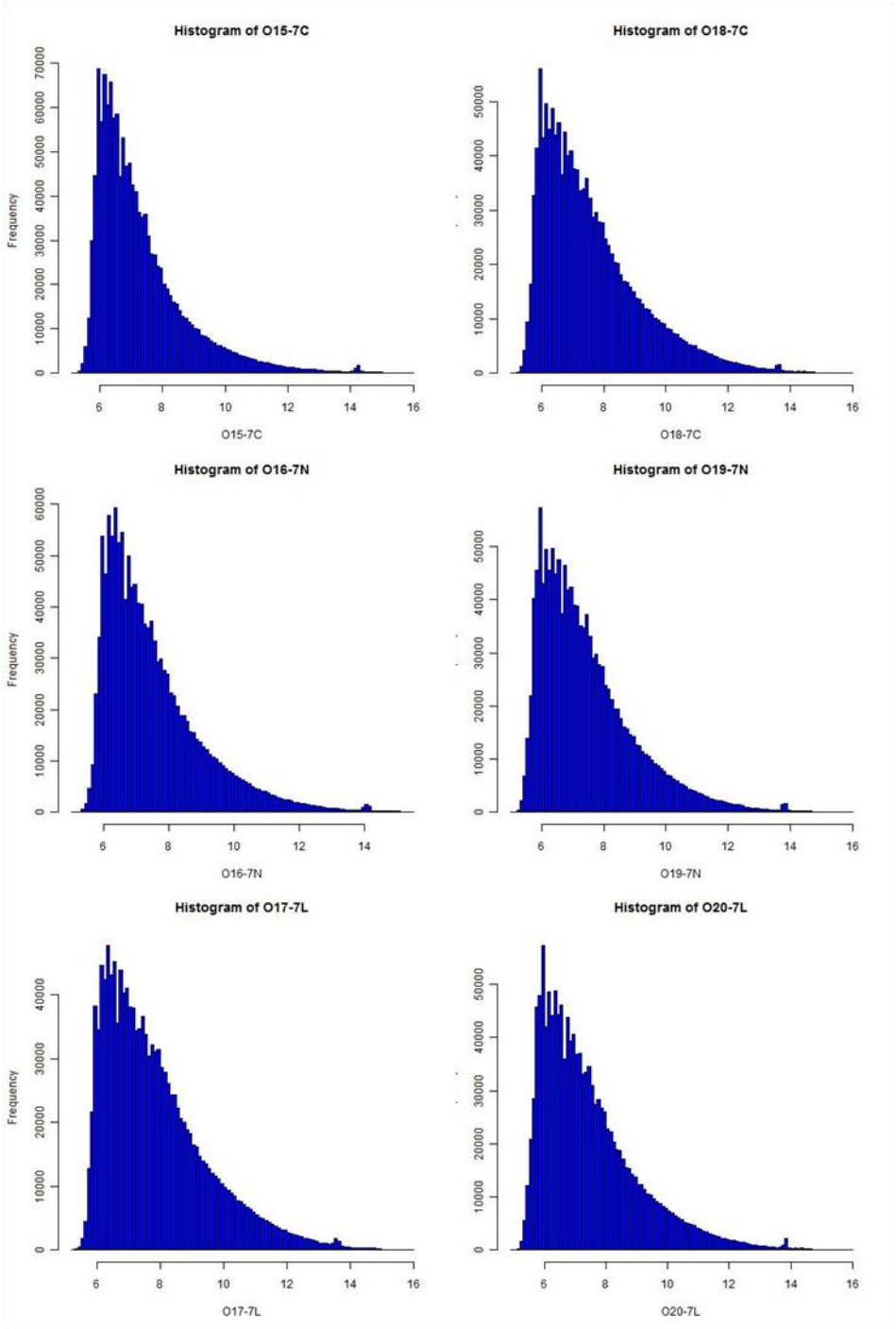

Figure 6: Frequencies of signal intensities along arrays. The plot shows the distribution of the signal intensity before normalization. Images were created with the help of `hist` function of the `graphics` package (R_Development_Core_Team, 2010).

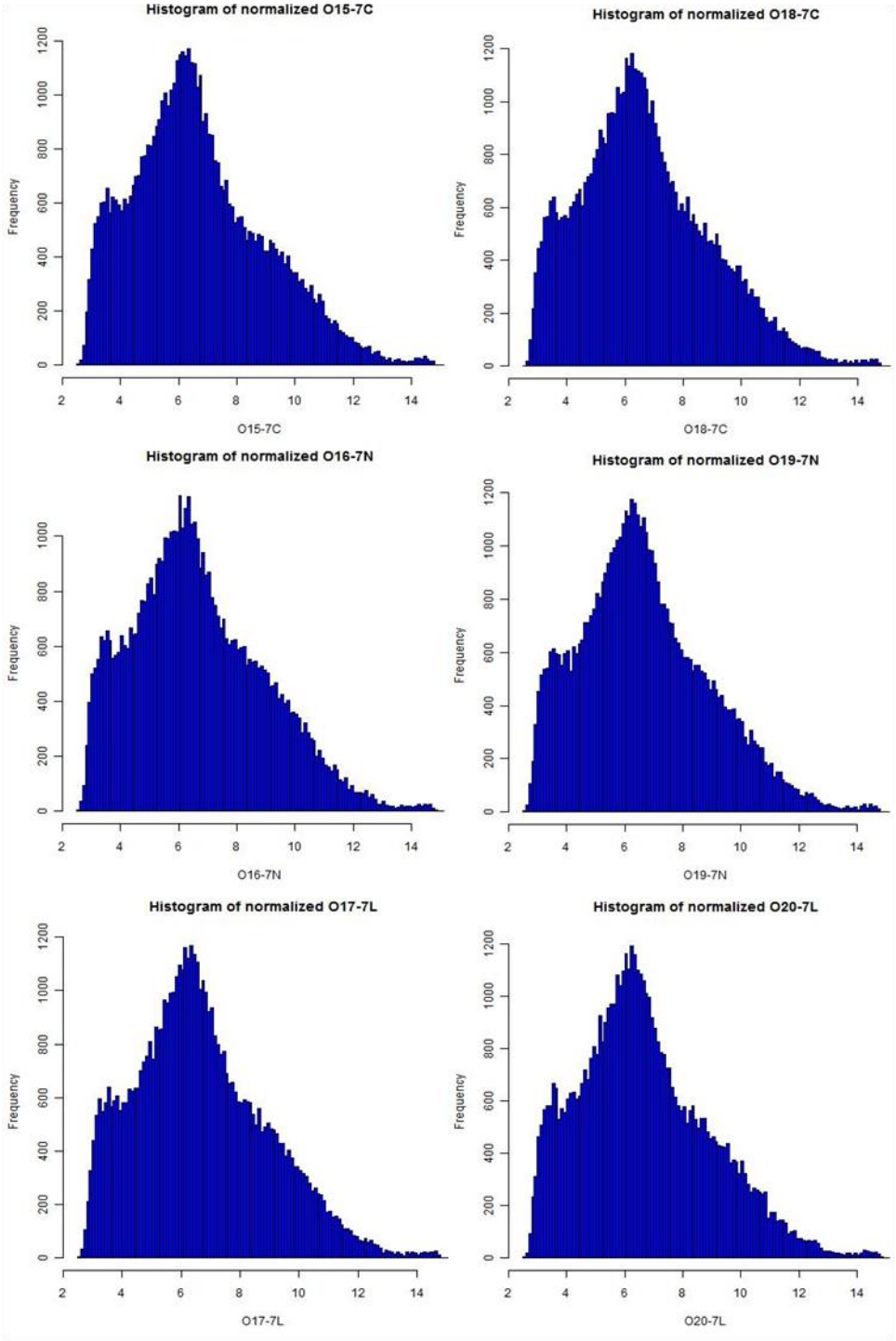

Figure 7: Frequencies of signal intensities along the arrays after normalization. The plot demonstrates the new distribution of the data following normalization. Normalization procedure simply approximates the data distribution to normal distribution on the grounds that following tests and analyses are based on the assumption that the data is distributed normally. Images were created with the help of `hist` function of the `graphics` package (R_Development_Core_Team, 2010).

.

Figure 9 shows the MA plots of both raw and normalized data indicating all arrays pass this quality control. The y-axis represents the 'M' component, intensity-ratio; x-axis represents the 'A' component, average ratio for a dot in the plot.

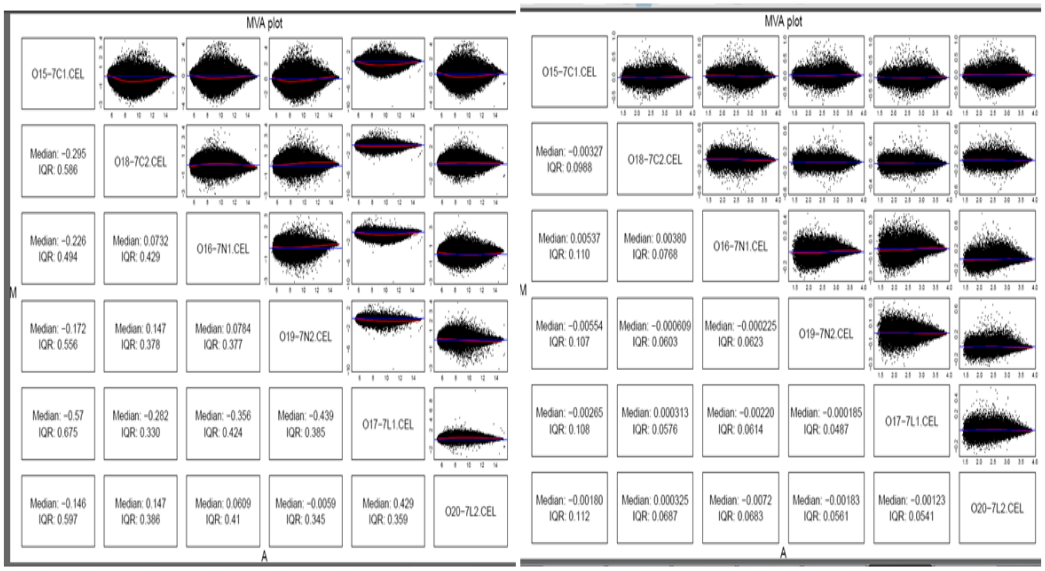

Figure 9 :MA Plots of arrays before and after normalization (from left to right). M (y-axis) is the intensity ratio; A (x-axis) is the average intensity for a dot in the plot. MA plot provides a quick overview of intensity-dependent ratio. Images were created with the help of `mva.pairs` function of the `affy` package (Gautier et al., 2004).

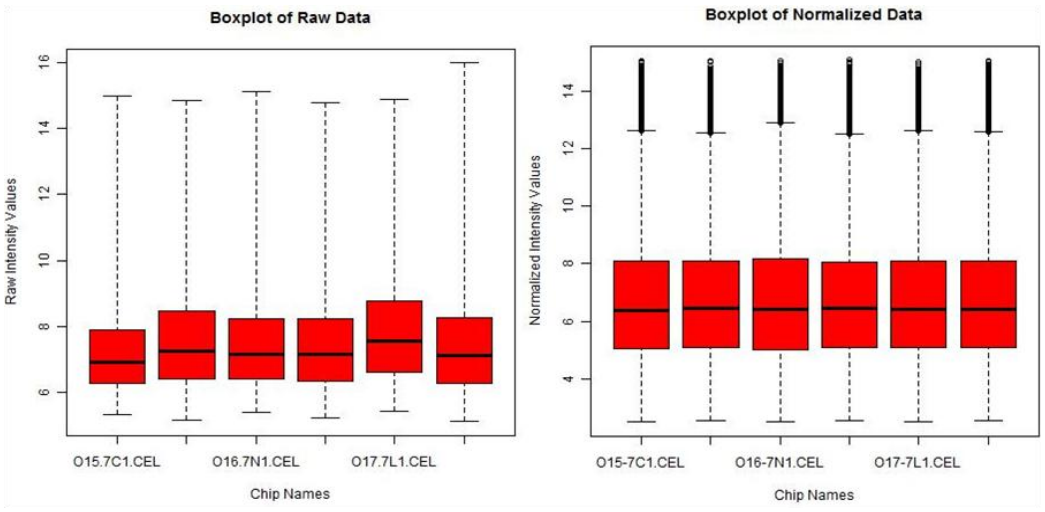

Figure 8: Box plot representation of data of each array before and after normalization. Represents intensity value distribution of arrays. To be analyzed together, all arrays should meet a very close mean value making the content comparable for differential expression analysis. y-axis shows signal intensity values (RMA normalization gives logged values of signal intensities). Images were created with the help of `boxplot` function of the `graphics` package (R_Development_Core_Team, 2010).

Figure 8 shows the boxplots of signal intensity distribution of arrays together pre- and post-normalization procedure.

We made use of `boxplot` function for generation of the figures. To make arrays comparable for statistical testing, they are expected to have similar distribution of signal intensity values after normalization. We used empirical Bayes moderated t-test with

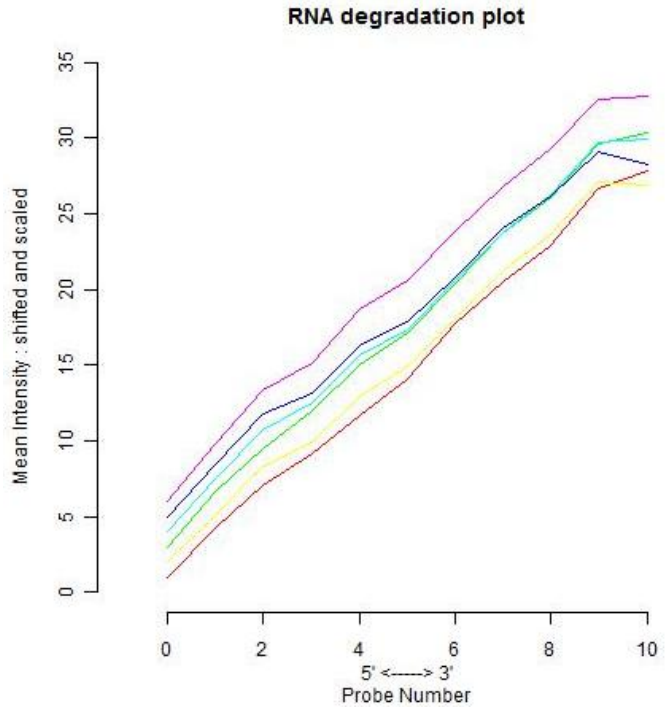

Figure 10: RNA degradation plot of arrays. Represents the mean intensity for 5' to 3' probes. Each line represents an array within the data. Agreement of slopes of all arrays indicates good quality and allows all arrays to be analyzed together. Image was generated with the help of `plotAffyRNAdeg` function of `affy` package (Gautier et al., 2004).

linear regression model. This type of moderated t-test is different from classical types of t-test, since with empirical Bayes method the probe-wise sample variances are shrunken towards a common value and the degrees of freedom for the individual variances are augmented (Smyth, 2004).

Next, RNA degradation profiles of the samples were subjected to quality control assessment, because high levels of RNA degradation lowers the hybridization and data quality, and biases the resulting expression profile. For this reason, we included generation of plots and tabularized reports for RNA degradation of arrays in the streamline routine. We used `AffyRNAdeg`, `summaryAffyRNAdeg`, `plotAffyRNAdeg` functions of `affyPLM` package.

Table 3: Summary of RNA degradation information of arrays. Note that slopes of all plots are comparable and below 3, the value advised by Affymetrix as a maximum. P value is calculated by `AffyRNAdeg` function internally from linear regression of means by number.

	O15-7C1.CEL	O18-7C2.CEL	O16-7N1.CEL	O19-7N2.CEL	O17-7L1.CEL	O20-7L2.CEL
slope	2.73	2.6	2.8	2.66	2.44	2.74
pvalue	1.23E-12	6.44E-11	3.36E-12	3.45E-11	1.31E-09	7.72E-11
RIN	7.7	7.8	7.7	8.1	7.7	7.9

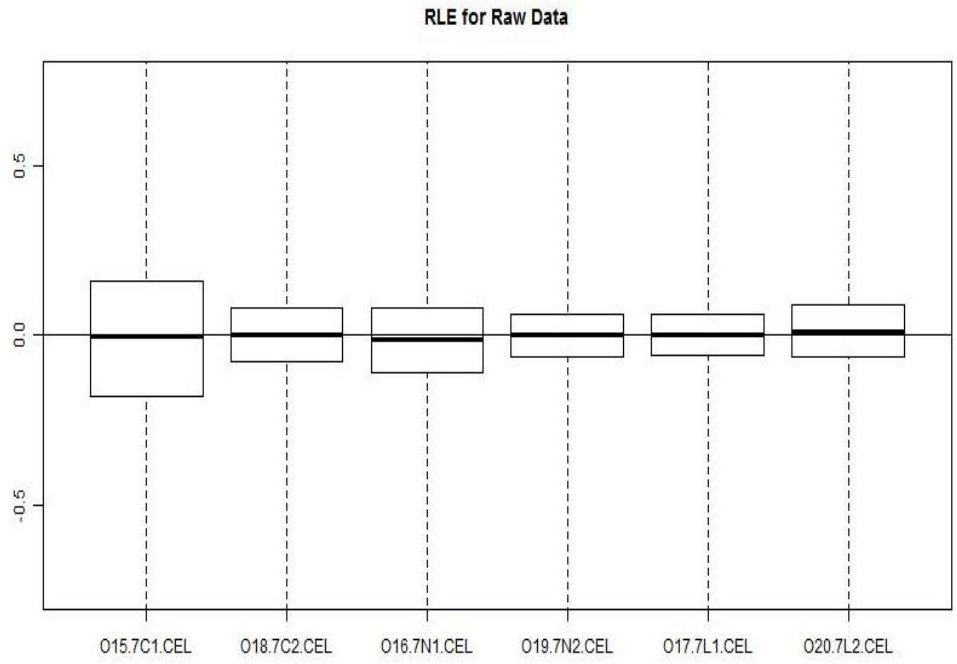

Figure 11: RLE (relative log expression) plots of arrays. Image was generated with the help of `RLE` function of the `affyPLM` package (Brettschneider et al., 2007).

Figure 10 shows RNA degradation plots of our six arrays. Ideally, without any RNA degradation, mean signal intensity would be flat from 5'-to-3' direction, as RNA degradation takes place in 5' to 3' direction, and signal from all portions would be equal. However, in reality the situation is not so perfect that it is always possible to see RNA degradation to some certain level. What is more important than the presence or absence of the degradation itself is the agreement of slopes between all arrays, which means all arrays have the same level of RNA degradation and all could be used together in normalization and analysis steps. Therefore, we were able to conclude that all of the six arrays pass this criterion for analysis and there exists no problem related with the qualities of signal data and normalization steps, as it is understood from post-normalization plots and RNA quality values and the RIN values (Table 3).

Finally, boxplots of the relative log expression, RLE plots (Moertel et al.) of the arrays was also generated by the routine for quality control assessment and arrays also passed that criterion as RLE plots of all have mean of zero (Figure 11). `RLE` function of the `affyPLM` package was used for this purpose.

4.2. IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

To determine differentially expressed genes, we used *DEGidentifier* function:

```
DEGidentifier(ndata,pdfile,lfc,method,adjust.method,p.value)
```

The function uses *normalized data* (ndata) and *experiment description file* (pdfile;

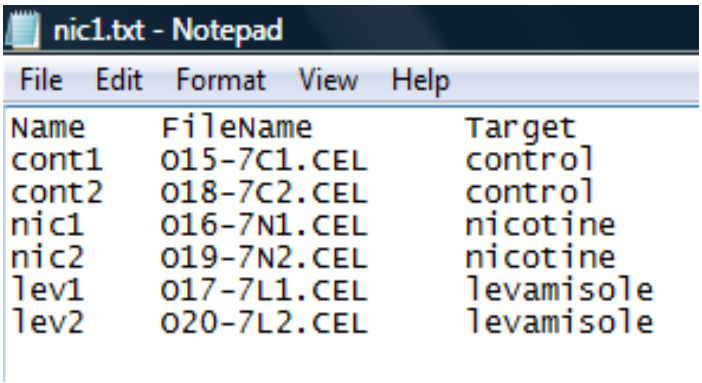

Figure 12: Experiment descriptor (phenodata) file we used for our analysis. We provide a screenshot of the tab delimited text file in order to demonstrate how a phenodata file should be provided for proper working of function *DEGidentifier*. Function reads in the .CEL files in the same order with column 'FileName' and makes pairs of which file belongs to which group in accordance with the order given in phenodata file. Since the group contrasts are made automatically by the routine according to the experiment description provided by the user, it is crucial to provide a properly prepared one to get reliable results.

column 'Target' used for groups and contrasts definition) as raw material to work on. In our analysis of levamisole and nicotine, an `lfc` (log fold change) threshold was not set not to overlook a systemic response and to obtain a certain number of probesets for functional analyses. We selected `method` (multiple testing method) as "separate", meaning that all contrasts were subjected to multiple testing separately, on the grounds that we tried here to answer independent questions by analyzing different contrasts

(please see “Limma User’s Guide” for details). For `adjust.method` (multiple testing correction method) we preferred “none”. There are a number of reasons for such preference. First and the foremost, multiple testing correction methods are applied to lower the possibility of Type I error (false-positive results) running the risk of losing true-positive results. When we applied Bonferonni (“BF”) or Benjamini&Hochberg (“BH”) correction methods we lose power because of small sample size. Therefore, we chose not to apply a multiple correction method but set a significance threshold relatively low, i.e., $p \leq 0.01$ (`p.value`). This value corresponded to a 27% FDR for nicotine-control comparison, and 26% FDR for levamisole-control comparison. Testing was performed following a linear model fitting and empirical Bayesian moderation. `lmFit`, `contrasts.fit`, `eBayes` and `decideTests` functions of `limma` package were used. Results were explained below.

4.2.1. NICOTINE AND LEVAMISOLE ACT SIMILARLY ON SW620 COLON CANCER CELLS

Overall, 3088 probes were found to be differentially expressed between any two groups ($p \leq 0.01$) (Figure 13). Nearly 2100 probesets showed differential expression in response

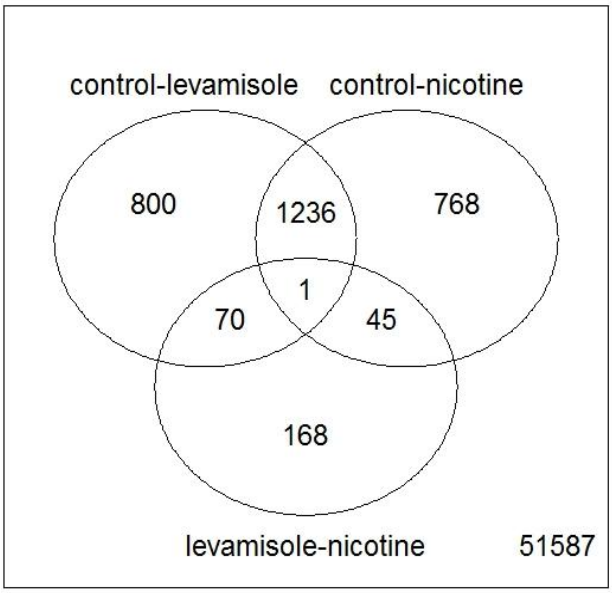

Figure 13: Venn Diagram representation of probe numbers with significant differential expression between the given groups (p≤0.01). Diagram obtained by the routine with `VennDiagram` function.

to levamisole and about the same number in response to nicotine. Interestingly, over 1200 of affected probesets were shared by nicotine and levamisole, and behaved in the same manner (i.e., up- and down-regulated in both) when compared to control samples (Figure 15 and Figure 16). On the other hand, about 200 probesets were affected differentially between the nicotine- and levamisole-treated groups. Accordingly, unsupervised hierarchical clustering showed that nicotine and levamisole groups were clustered together, while contrasting with the control group samples (Figure 14). When the genes with differential expression were grouped in terms of the direction of the response, it was obvious that about 60% of the affected genes responded with up-regulation in both levamisole and nicotine treatment as compared to control group.

Table 4: Table showing the number of probes differentially expressed between groups. (*) Gene numbers are obtained by ID conversion. Probe IDs with no known gene symbols (NAs) are removed from gene lists.

Comparison	Up-regulated Probe/Gene*	Down-regulated Probe/Gene*	Sum Probe/Gene*
levamisole-control	1275/837	832/558	2107/1395
nicotine-control	1297/833	753/511	2050/1344
nicotine-levamisole	254/153	30/22	284/175

Therefore it might be possible to suggest that chronic nicotine or levamisole exposure on the SW620 cells had rather up-regulatory role than down-regulatory.

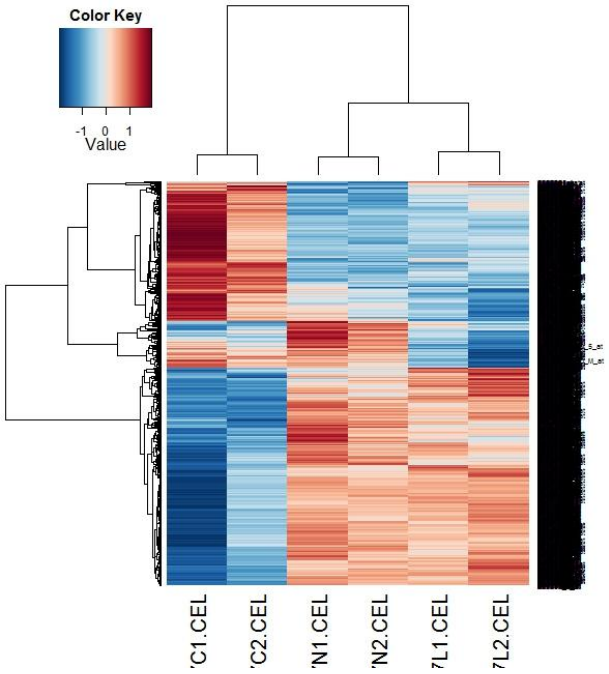

Figure 14: Heatmap generated with the expression values of significantly differentially expressed genes. Note the similarity between the levamisole- and nicotine-treated samples (p<=0.01) while both are contrasting with control samples. Red color: up-regulation; Blue color: down-regulation. Picture generated by the routine using `heatplot` function of the package `made4` (Culhane).

A few group of genes, differentiating nicotine response from levamisole response, were largely (87-89%) subjected to up-regulation by nicotine treatment (Table 4). The percentages given also were similar when probe-to-gene conversion was made.

Table 5: Detailed numbers of probes differentially expressed between groups.

	levamisole-control	nicotine-control	nicotine-levamisole
2+ fold UP	11	18	
1-2 fold UP	248	316	46
0-1 fold UP	1016	963	208
0-1 fold DOWN	560	530	30
1-2 fold DOWN	246	194	
2-3 fold DOWN	16	18	
3+ fold DOWN	10	11	

Detailed numbers of differentially expressed probes distributed to the difference clusters were given in Table 5. Interestingly, we found that moderate but not high levels of fold changes were seen in levamisole and in nicotine treatment. In both treatments about 75% of the differentially expressed probes responded in the range of *0-to-1* fold change in both directions.

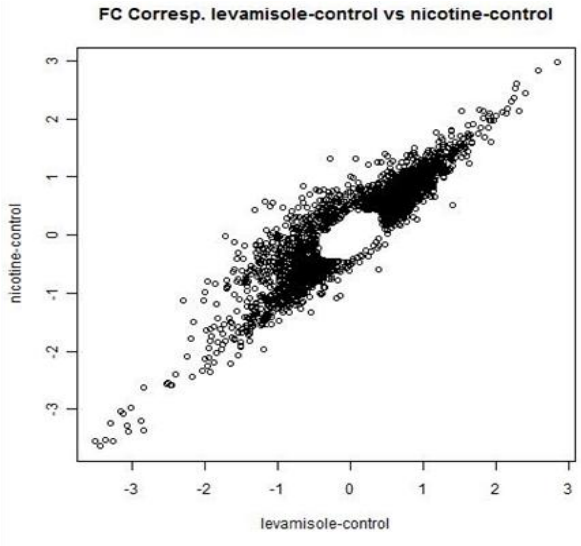

Figure 15: Fold change correspondence between gene groups affected by nicotine and levamisole in comparison with control. Image was generated with the help of `plot` function of the `graphics` package (R_Development_Core_Team, 2010).

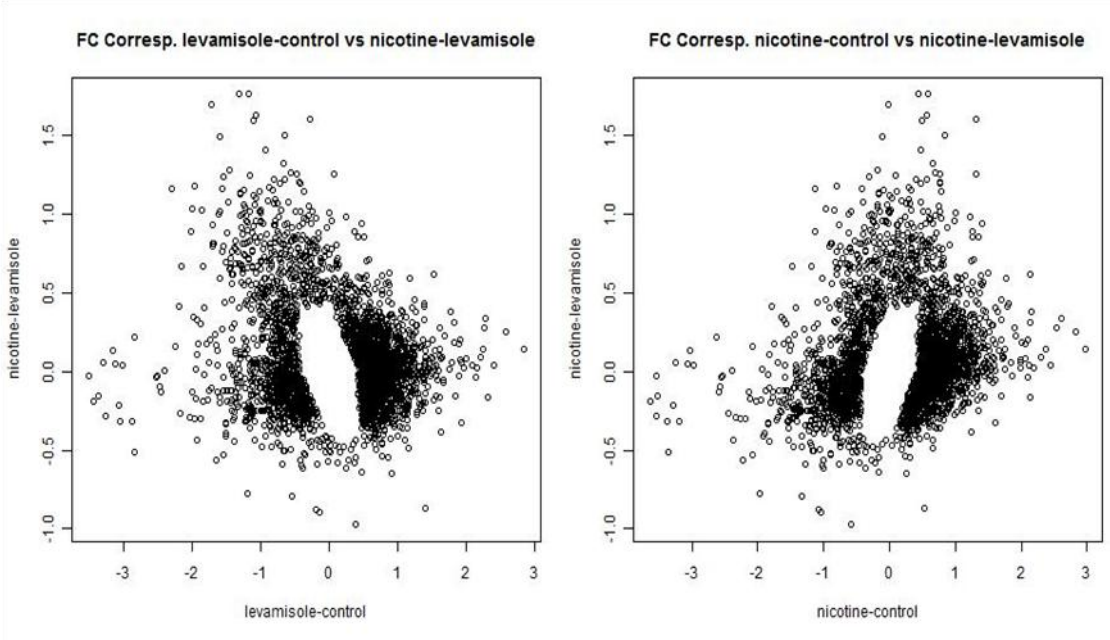

Figure 16: Fold change correspondence between gene groups affected by either levamisole (left) or by nicotine versus gene groups that differ between nicotine and levamisole. Image was generated with the help of `plot` function of the `graphics` package (R_Development_Core_Team, 2010).

4.3. FUNCTIONAL ANALYSIS

Functional analyses of the data were performed with the help of *FuncAnn* module. This third module is composed of six different functions: *GO_HyperG*, *KG_HyperG*, *KEGGscript*, *PlotCharts*, *ChartPlotter* and *Grapher*.

```
GO_HyperG("BP", conditional=TRUE, AllSignifProbes,
universe_GeneIds=NULL,categorySubsetIds=NULL,
"hgu133plus2.db", 0.05, "over")

KG_HyperG(AllSignifProbes, universe_GeneIds = NULL,
"hgu133plus2.db","KEGG", 0.05, "over")
```

`GO_Hyper` and `KEGG_Hyper` are designed to perform hypergeometric tests for over- or under-representation of GO terms and KEGG pathways, respectively. Results are saved to the working directory in spreadsheet or plain text formats.

```
KEGGscript(results)
```

KEGGscript maps all the significant DEGs to the available pathways and provides a detailed result file informing which of significant DEGs are mapped to any KEGG pathway and at what percentage. Although `KG_HyperG` performs a similar job with a user-defined p-value, *KEGGscript* provides virtually all pathway information of the results. We believe such information would be useful particularly for the assessment of a systemic effect by the treatment. When numbers of genes in a given gene list are represented in a pathway with low percentage, it is eliminated in complete pathway testing, because such gene testing approach will require as many genes in the pathway. However, if the treatment affects many pathways, nonetheless with low coverage percentages, it could be considered for a systemic effect. Therefore, for a more systems biology geared approach, *KEGGscript* would be beneficial.

PlotCharts is the subfunction of *ChartPlotter* and the latter is employed in *KEGGscript* to plot graphical representation of distribution of top five GO and KEGG categories in significant gene lists. This plotting process were performed for each significant gene list of contrasts as well as for list of all significant DEGs in the data.

Grapher generates the KEGG pathway graphs for all contrasts with color coded up- or down-regulation information and directed edges. This function has been designed to take three inputs: normalized data, fold-change data of significant DEGs and finally a vector of KEGG ids desired to be plotted. Plotting all of the KEGG ids found in the data is

unnecessary and brings much burden to the CPU. Therefore, we chose to generate graphs of KEGG pathways that were previously found to be significant in the KEGG Hypergeometric Testing and retained this as default list if the `SignifKEGGIDs` input is given as NULL.

Table 6: Hypergeometric testing results of GO terms of genes differentially expressed in response to nicotine. Table is generated by *GO_HyperG* function. There were about 200 significantly overrepresented GO terms (see Appendix). Top 15 are provided here (p≤0.05).

GOBPID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0006695	1.88E-05	6.137	2.507176674	11	33	cholesterol biosynthetic process
GO:0006066	2.77E-05	2.389935846	15.50045893	33	207	alcohol metabolic process
GO:0006694	8.14E-05	2.969618129	7.825430225	20	103	steroid biosynthetic process
GO:0055114	0.00012371	1.673508435	45.73698054	71	602	oxidation reduction
GO:0008654	0.000143128	2.924697581	7.521530022	19	99	phospholipid biosynthetic process
GO:0043603	0.000186888	3.800505773	4.17862779	13	55	cellular amide metabolic process
GO:0045017	0.000333598	3.025836877	6.153979109	16	81	glycerolipid biosynthetic process
GO:0005996	0.000420841	2.049759986	17.09438641	32	225	monosaccharide metabolic process
GO:0045007	0.000437342	Inf	0.227925152	3	3	depurination
GO:0051186	0.000838224	2.094512195	14.13135943	27	186	cofactor metabolic process
GO:0006006	0.001095333	2.199749988	11.52882808	23	152	glucose metabolic process
GO:0048255	0.001130176	6.667571235	1.291575862	6	17	mRNA stabilization
GO:0019320	0.001609768	2.899125524	5.166303449	13	68	hexose catabolic process
GO:0009820	0.001685911	3.499643214	3.418877283	10	45	alkaloid metabolic process
GO:0001568	0.00175715	1.78639481	21.65288946	36	285	blood vessel development

GO term hypergeometric testing was performed with significant probes lists for nicotine and levamisole treatments with the help of *GO_HyperG* function. There was again a large similarity between nicotine and levamisole since both groups shared 141 of 191 and 150 passed GO terms, respectively. Table 6 and Table 7 indicate the first 15 GO IDs of nicotine and levamisole. Similarities between most significant list can be listed as cholesterol biosynthetic, steroid biosynthetic and alcohol metabolic processes.

Table 7: Hypergeometric testing results of GO terms of genes differentially expressed in response to levamisole. Table is generated by *GO_HyperG* function. There were about 150 significantly overrepresented GO terms (see Appendix). Top 15 are provided here. p≤0.05.

GOBPID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0048519	0.000140646	1.378847354	144.3841587	185	1779	negative regulation of biological process
GO:0006695	0.000194961	4.958959772	2.678289622	10	33	cholesterol biosynthetic process
GO:0006986	0.000357035	3.332277986	5.031938078	14	62	response to unfolded protein
GO:0031324	0.000543165	1.527808823	57.29916585	82	706	negative regulation of cellular metabolic process
GO:0048522	0.000706006	1.329330832	144.3841587	180	1779	positive regulation of cellular process
GO:0016481	0.000832369	1.64242941	36.52213121	56	450	negative regulation of transcription
GO:0051172	0.001285224	1.573932339	41.31058841	61	509	negative regulation of nitrogen compound metabolic process
GO:0006613	0.001702696	8.119601329	0.973923499	5	12	cotranslational protein targeting to membrane
GO:0032535	0.002056754	1.702379753	25.80897272	41	318	regulation of cellular component size
GO:0010941	0.002345488	1.418551302	66.30795822	89	817	regulation of cell death
GO:0051270	0.0032171	1.886022514	15.50161569	27	191	regulation of cellular component movement
GO:0044282	0.00390167	1.790922619	18.01758473	30	222	small molecule catabolic process
GO:0030968	0.004049797	4.873503591	1.623205832	6	20	endoplasmic reticulum unfolded protein response
GO:0006066	0.004052468	1.805127235	17.28940529	29	215	alcohol metabolic process
GO:0000122	0.004468916	1.720200573	20.53355377	33	253	negative regulation of transcription from RNA polymerase II promoter

After we assessed the molecular and cellular processes affected by nicotine or levamisole treatments, we evaluated the specific reaction and/or gene interaction information with KEGG pathways in order to assess functional effects of the treatments in detail. We performed KEGG pathway hypergeometric testing with DEG of "nicotine vs. control" and "levamisole vs. control" comparisons. Each result included 13 different pathways and 11 of which were the same (Table 8). The most important part for us was molecular processes and pathways involved in steroid, lipid and cholesterol metabolisms as this finding was previously not shown in the literature to our knowledge.

Table 8: KEGG Pathways overrepresented in nicotine and levamisole data. Lists generated by *KG_HyperG* function (p≤0.05).

Nicotine	Levamisole	Nicotine-Levamisole
Glycolysis / Gluconeogenesis	Glycolysis / Gluconeogenesis	Galactose metabolism
Citrate cycle (TCA cycle)	Citrate cycle (TCA cycle)	Steroid biosynthesis
Pentose phosphate pathway	Pentose phosphate pathway	Oxidative phosphorylation
Pentose and glucuronate interconversions	Pentose and glucuronate interconversions	Purine metabolism
Fructose and mannose metabolism	Fructose and mannose metabolism	Valine, leucine and isoleucine degradation
Galactose metabolism	Galactose metabolism	Starch and sucrose metabolism
Ascorbate and aldarate metabolism	Ascorbate and aldarate metabolism	Amino sugar and nucleotide sugar metabolism
Fatty acid biosynthesis	Fatty acid biosynthesis	Inositol phosphate metabolism
Fatty acid metabolism	Fatty acid elongation in mitochondria	One carbon pool by folate
Synthesis and degradation of ketone bodies	Fatty acid metabolism	Riboflavin metabolism
Steroid biosynthesis	Synthesis and degradation of ketone bodies	Nicotinate and nicotinamide metabolism
Primary bile acid biosynthesis	Steroid biosynthesis	Pantothenate and CoA biosynthesis
Ubiquinone and other terpenoid-quinone biosynthesis	Primary bile acid biosynthesis	Aminoacyl-tRNA biosynthesis
		Metabolic pathways
		RNA degradation
		Basal transcription factors
		Spliceosome
		Proteasome
		Protein export
		Nucleotide excision repair
		MAPK signaling pathway

Although the significant number of DEGs in the levamisole-nicotine comparison was smaller than the other two, over represented pathway number was much higher. Such results might arise because of testing of small gene lists. Nevertheless, keeping the p-value lower can provide stringency with such small gene lists.

As a third step in functional analysis part, we employed *KEGGscript* function providing detailed information in both a tabular and graphics format. *KEGGscript* uses all the KEGG information available from the data in these analyses (Table 8). According to these results about 60% of the genes in the steroid biosynthesis pathway were up-

regulated with nicotine and about 50% was up-regulated with levamisole treatment. Other most affected pathways were: Terpenoid backbone biosynthesis, Limonene and pinene degradation, Propanoate metabolism, Valine, leucine and isoleucine degradation, beta-Alanine metabolism, Fatty acid biosynthesis Fatty acid metabolism, Biosynthesis of unsaturated fatty acids and synthesis and degradation of ketone bodies.

Table 9: KEGG pathways effected by nicotine or levamisole treatment. All available KEGG pathways are evaluated in terms of genes and pathway coverage percentage values. Top results are provided here; for full result see Appendix. Table was generated by *KEGGscript* function.

GROUPS	Levamisole-Control		Nicotine-Control		Nicotine-Levamisole	
PATHWAY NAME	**DOWN %**	**UP %**	**DOWN %**	**UP %**	**DOWN %**	**UP %**
Steroid biosynthesis	-	47.06	-	58.82	-	5.88
Terpenoid backbone biosynthesis	-	26.67	-	26.67	-	-
Limonene and pinene degradation	-	25.00	-	25.00	-	-
Propanoate metabolism	3.03	24.24	-	24.24	-	-
Valine, leucine and isoleucine degradation	-	22.73	-	20.45	-	2.27
beta-Alanine metabolism	4.55	22.73	4.55	9.09	-	-
One carbon pool by folate	5.88	17.65	-	5.88	-	5.88
Fatty acid biosynthesis	-	16.67	-	16.67	-	-
Fatty acid metabolism	-	16.67	-	14.29	-	-
Pyruvate metabolism	2.50	15.00	-	17.50	-	-
Pentose phosphate pathway	-	14.81	-	18.52	-	-
Glycerolipid metabolism	4.08	14.29	4.08	18.37	-	-
Ubiquinone and other terpenoid-quinone biosynthesis	-	14.29	-	14.29	-	-
Butanoate metabolism	-	14.29	-	11.43	-	-
Biosynthesis of unsaturated fatty acids	4.76	14.29	-	9.52	-	-
Thiamine metabolism	-	12.50	12.50	12.50	-	-
Synthesis and degradation of ketone bodies	-	11.11	-	11.11	-	-

Names of the genes were provided in "***KEGG_RESULT.txt" files generated by *KEGGscript* for each comparison. Pathway coverage percentages were provided in spreadsheet files with "KEGG_TABLE***" denomination.

In addition to these, *KEGGscript* also provides graphics representation of top 5 categories for both GO terms and KEGG pathways for the significant DEGs of each group and significant DEGs in all data. The number of genes covered, rather than the coverage percentage of a GO term pathway, is the determinant of ranking in these plots. For this visual support we made use of `GeneAnswers` package components and we modified two of functions, plot functions, in order to adjust them in the customized routine used in the present analysis.

Figure 17 to Figure 21 examplified the GO and KEGG plots for upregulated and downregulated genes in nicotine and levamisole treatment, respectively. Similarity betwen the levamisole and nictine response was apparent in particularly the up-regulatory sections. The most affected 5 GO terms and KEGG pathways by levamisole or nicotine were nearly the same. Shared GO processes were metabolic processes, small molecule metabolic process, lipid metabolic process and oxidation reduction processes, all of which were up-regulated by both nicotine and levamisole. KEGG pathways common to these treatments were metabolic pathways, valine, leucine and isoleucine degradation, steroid biosynthesis and propanoate metabolism. Fifth pathway in nicotine vs control group was the glutathione metabolism pathway while it was the fatty acid metabolism in the levamisole treatment.

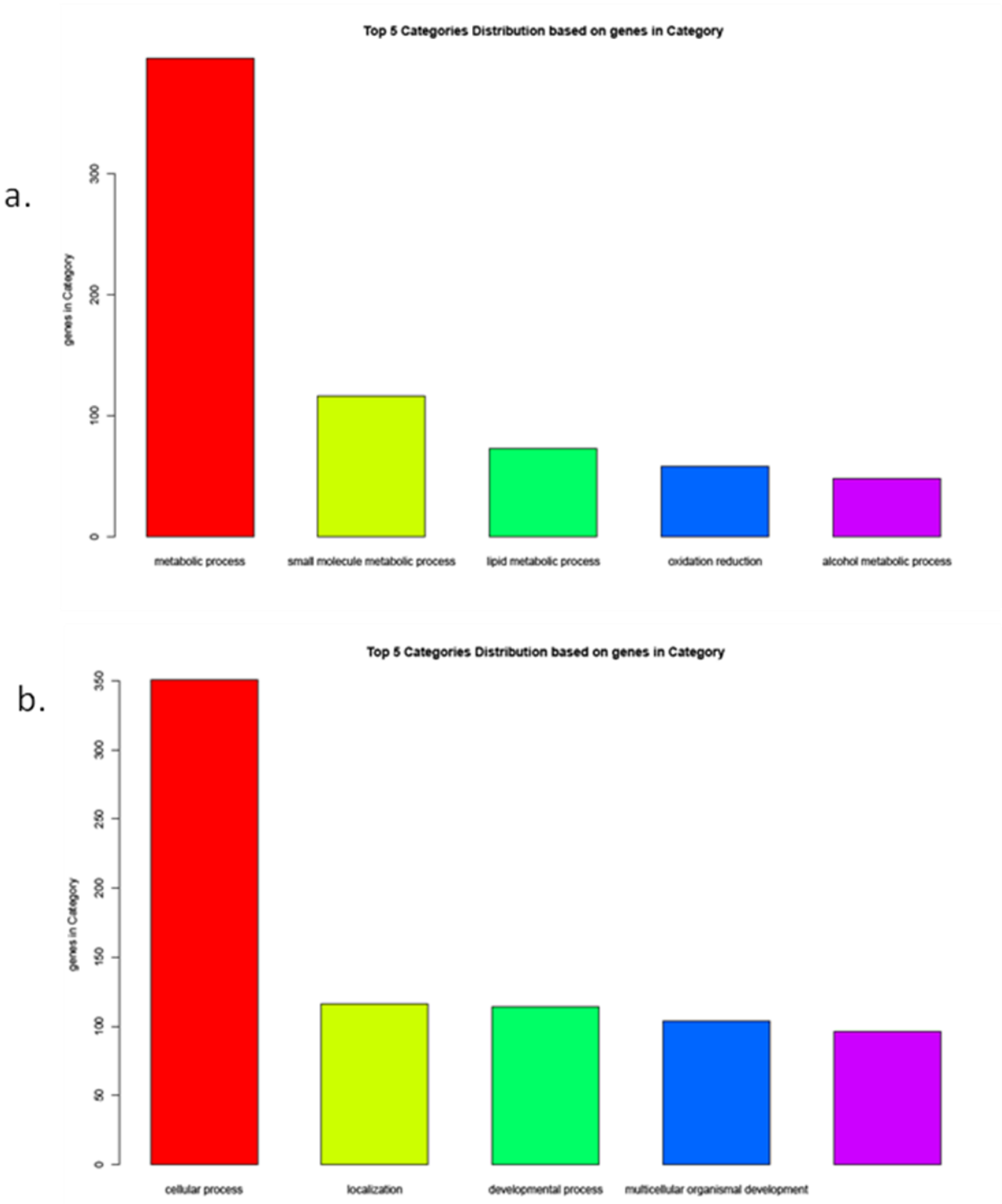

Figure 17: Bar plots of top 5 categories in GO Biological Processes of DEGs up- or down-regulated by nicotine. Generated by KEGGscript function. ChartPlotter and PlotCharts functions are used for plotting. Function benefits from GeneAnswers package of BioConductor. a)GO Processes in up-regulated DEGs by nicotine; b) GO Processes in down-regulated DEGs by nicotine. Images were generated by *KEGGscript* function with the help of *ChartPlotter* and *PlotCharts*, two functions modified from equivalent functions of `GeneAnswers` package (Feng et al., 2001).

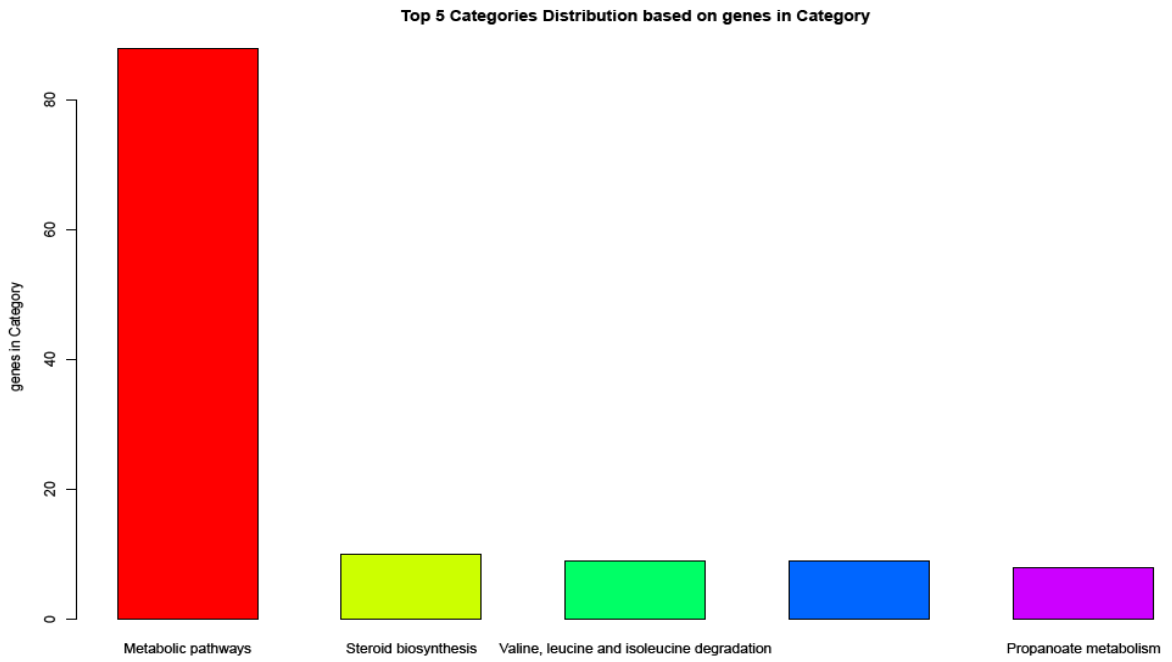

Figure 19: Barplots of top 5 categories in KEGG Pathways of DEGs down-regulated by nicotine. Images was generated by *KEGGscript* function with the help of *ChartPlotter* and *PlotCharts*, two functions modified from equivalent functions of `GeneAnswers` package (Feng et al., 2001).

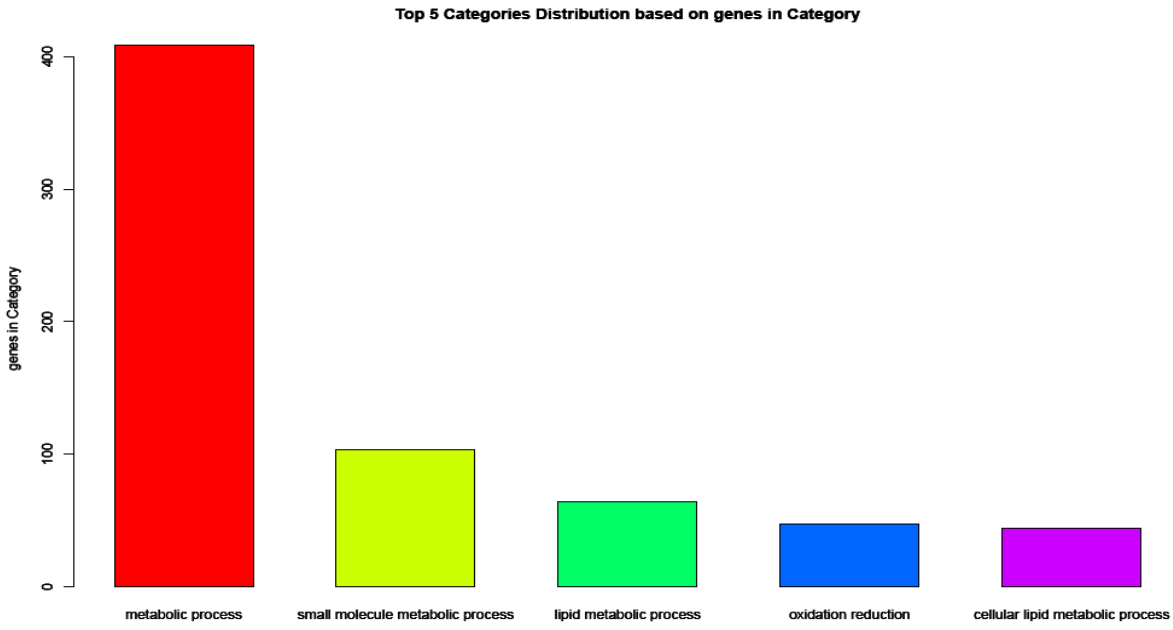

Figure 18: GO Biological Processes of DEGs up-regulated by levamisole. Images was generated by *KEGGscript* function with the help of *ChartPlotter* and *PlotCharts*, two functions modified from equivalent functions of `GeneAnswers` package (Feng et al., 2001).

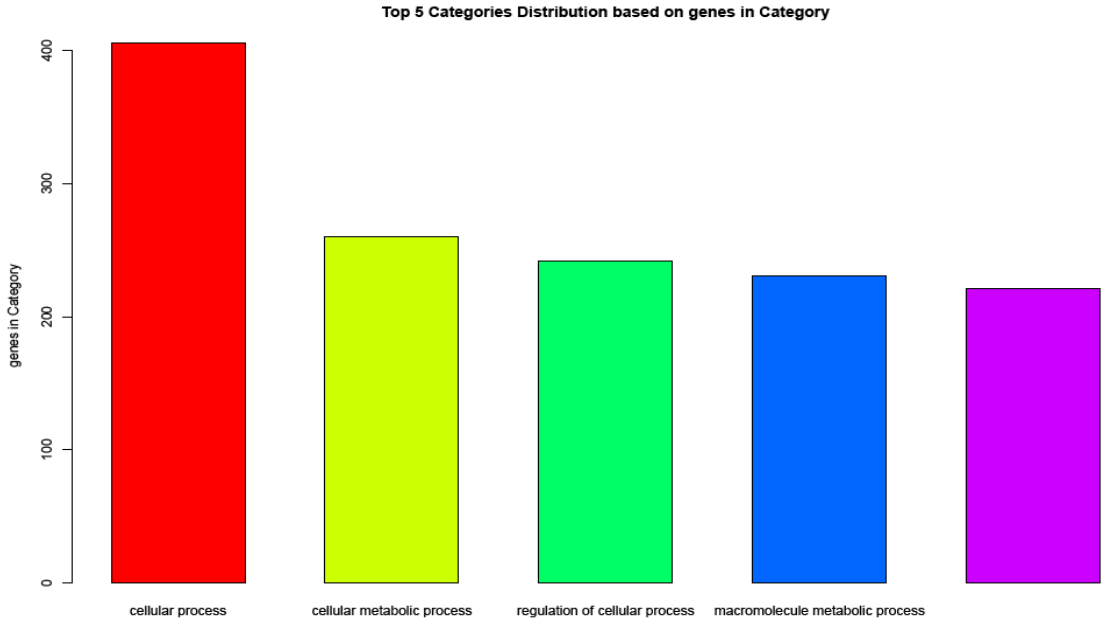

Figure 21: GO Biological Processes of DEGs down-regulated by levamisole. Images was generated by *KEGGscript* function with the help of *ChartPlotter* and *PlotCharts*, two functions modified from equivalent functions of `GeneAnswers` package (Feng et al., 2001).

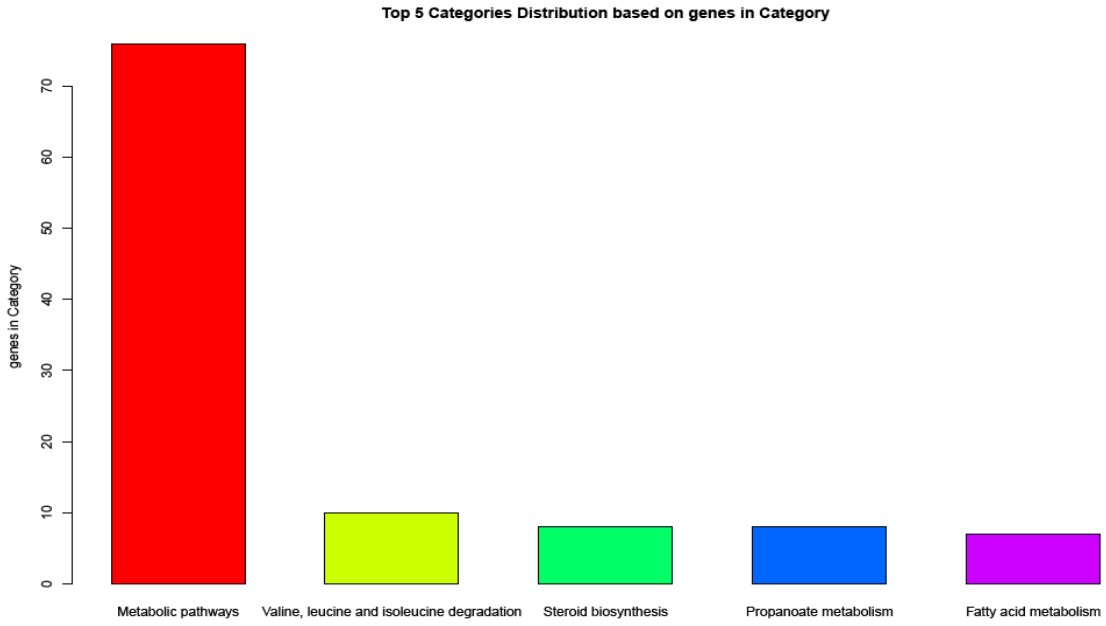

Figure 20: KEGG Pathways in DEGs up-regulated by levamisole. Images was generated by *KEGGscript* function with the help of *ChartPlotter* and *PlotCharts*, two functions modified from equivalent functions of `GeneAnswers` package (Feng et al., 2001).

Finally, *Grapher* function provides KEGG pathway graphics with up- or down-regulation color code and directed edges. Basicly, `KEGGgraph` objects of pathways are visualized with this function. However, definition of attributes for color code and log-fold change data is defined by via *Grapher* function so that the graph only becomes informative when plotted in an experimental context. Pathways found significant from *KG_HyperG* function in any of the comparisons are plotted.

The upregulated genes in the steroid biosyntheis pathway for all three comparisons were shown in Figure 24. This pathway was chosen as a representative since it has been one of the most significantly affected pathways. Pivotal enzymes in cholesterol synthesis from dehydrocholesterols, DHCR7 and DHCR24, and lanosterol synthase (LSS) catalyzing the first step in biosynthesis of cholesterol, steroid hormones and vitamin D, were significantly upregulated in both treatments (Figure 24).

4.4. CONFIRMATORY REALTIME PCR RESULTS

We performed confirmatory real time quantitative PCR analyses with chronic (7 days) nicotine and levamisole treated SW620 cells for *GULP1*, *CHRNA1*, *CHRNB1*, *MKI67*, *XIAP*, *FOSL2* and *MAP1B*, genes whose transcriptional responses were summarized in Table 10. We have selected these genes for confirmation experiments since these genes except CHRNB1 also seemed to be affected by 1μM nicotine treatment in serum deprived SW620 cells in a previous study by our group (Kaya, 2009). Furthermore, we used additional set of samples with treatments of nicotine and levamisole for 15 days to see the effects associated with chronic use. The real time qPCR analyses for 7 and 15 days were performed using an independent cell culture experiments; each day had a

single replicate, thus are preliminary in nature. Figure 22 shows the efficiency curves of the genes under investigation; the primer specific efficiencies were used in calculations of fold changes between control and treatment groups.

Except the proliferation marker *MKI67*, we were able to confirm responses of all the selected genes to the treatments (Figure 23; Table 10). Accordingly, *GULP1* was drastically reduced as *FOSL2*, and *MAP1B*. *XIAP* was moderately reduced at day 7 (Table 10). The fold change differences between the control and treatment groups in microarray experiments and realtime qRT PCR results from 7 and 15 day experiments were shown in Figure 23. The day 7 qRTPCR experiment was in concordance with day 7 microarray results, as expected. At day 15, the changes observed at day 7 might be reversed except for CHRNA1 and CHRNB1. These results were obtained from one set of samples thus need to be repeated in future studies.

Table 10: Fold changes of confirmation genes in microarray data and qPCR experiment. qPCR fold change values were calculated with ΔΔCt method (Pfaffl, 2001) and original results were logged. Note the similarity between profile pairs, except for MKI67. Correlation analyses between profiles from microarray fold-change and qPCR fold-change were performed: for nicotine ρ = 0.9175, p-value = 0.0099; for levamisole: ρ = 0.9652, p-value = 0.0018. (MKI67 was excluded from correlation analyses).

	Microarray Results		**qPCR results**		
	1 µM Nic	**1 µM Lev**	**1 µM Nic**	**1 µM Lev**	**10 µM Lev**
GULP1	-3.25	-3.09	-1.17	-1.04	-1.50
CHRNA1	0.75	0.63	0.80	0.45	0.38
CHRNB1	0.66	0.48	0.33	0.36	0.45
MKI67	0.69	0.66	-0.35	-0.12	-0.04
XIAP	-0.10	-0.58	-0.14	-0.28	-0.49
FOSL2	-1.03	-0.96	-0.53	-0.54	-0.77
MAP1B	-0.60	-1.18	-0.58	-0.36	-0.31

CHRNA1 and *CHRNB1*, encoding two subunits of nAChRs involved in the congenital myasthenic syndrome, increased in expression in response to both of the treatments in the PCR experiments, in line with the microarray findings. Interestingly, no other

subunit expression difference was seen either in PCR experiments or microarray data, since it was previously reported that nicotine exposure had up-regulated nAChR subunit expression (Fu et al., 2009).

XIAP, known to encode an anti-apoptotic gene and mediator of caspase-3 and -7 (Scott et al., 2005), decreased with chronic levamisole and nicotine treatments in the microarray data (Table 10) and this was confirmed with PCR experiments. Nicotine's anti-apoptotic ability by inducing *XIAP* has been shown in a previous study (Dasgupta et al., 2006). However, that study was not based on a chronic treatment and they used 36h samples.

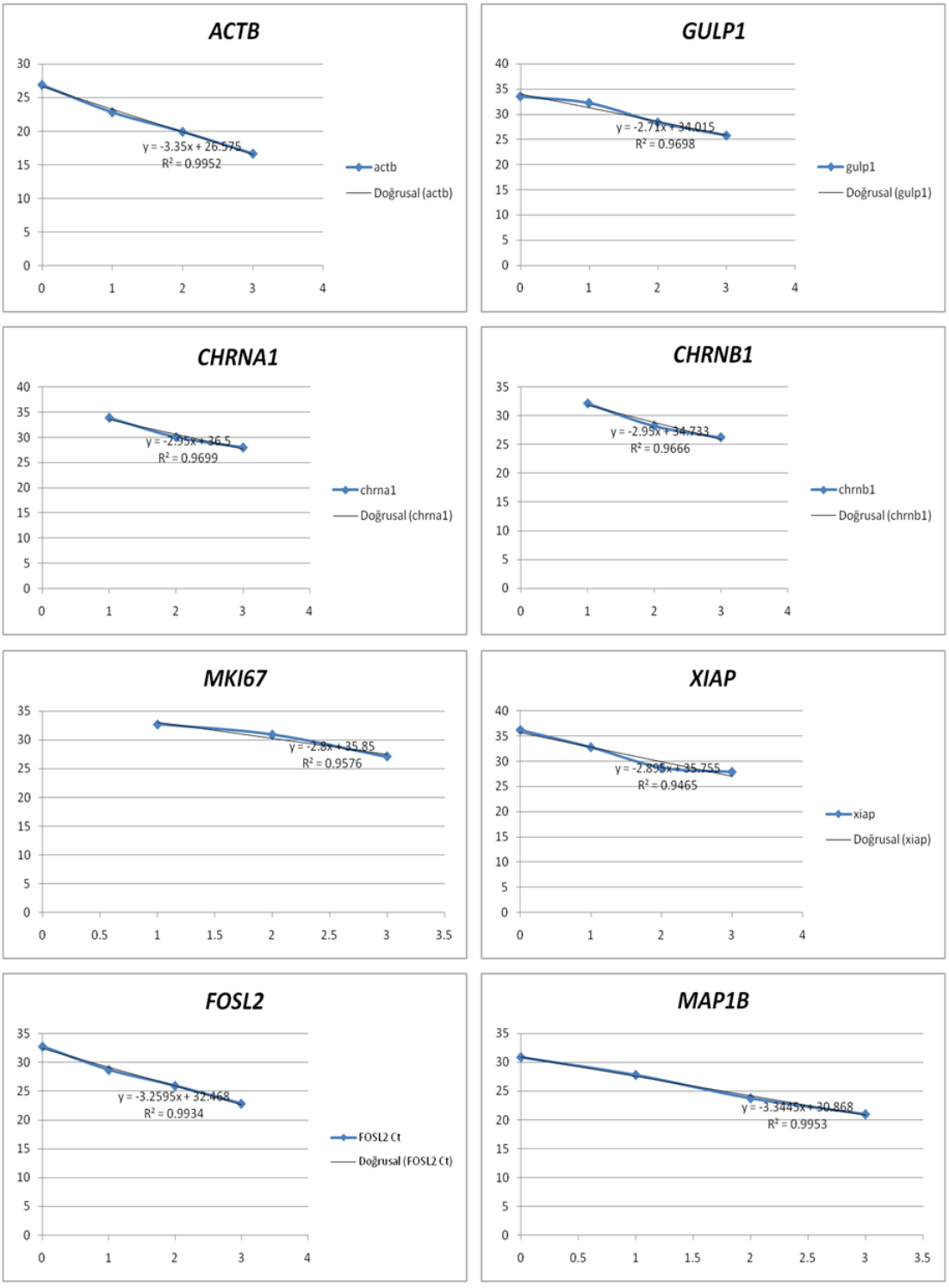

Figure 22: Efficiency curves for primers of the selected confirmation genes. The x-axis shows logged values of relative concentration points of the cDNA samples (1000 for undiluted standard cDNA; 100 for 1:10 dilution; 10 for 1:100 dilution and 1 for 1:10000 dilution) and the y-axis shows the Ct values.

Decreases in *FOSL2* and *MAP1B* were also confirmed with real-time PCR experiments. *FOSL2* (FOS-like antigen 2; or Fra-2: Fos-related antigen 2), encodes a member of FOS family of proteins and was implicated in cell proliferation and differentiation an upstream regulator of AP-1 (Eferl et al., 2008).

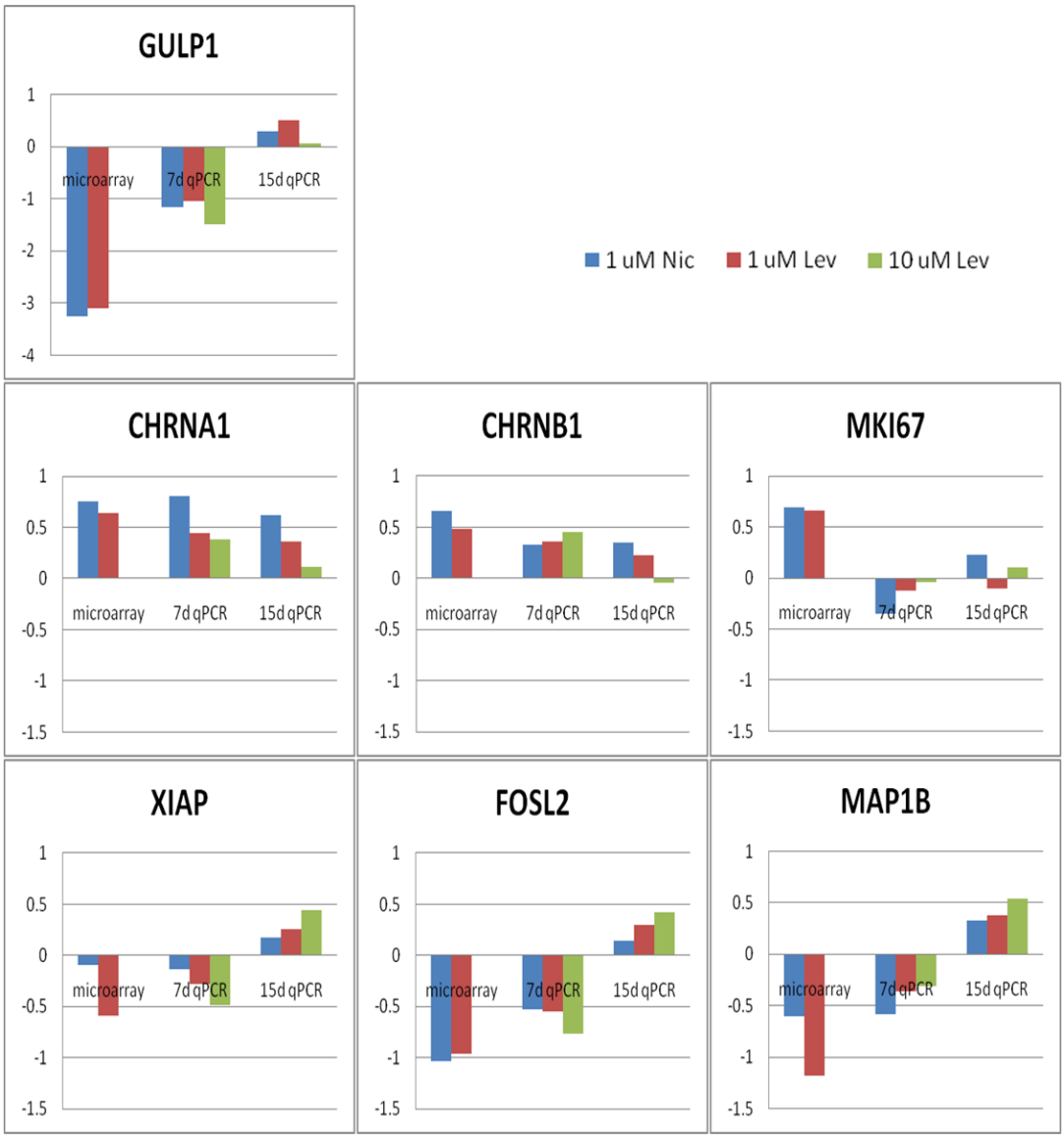

Figure 23: Confirmatory real-time results of given genes. All except MKI67 are in line with the microarray data findings (see Table 10 for correlation analyses results).

MAP1B, encoding a microtubule associated protein and known neuronal marker (Schoenfeld et al., 1989), was down-regulated upon chronic nicotine and levamisole exposure in both the microarray and 7 day qRT-PCR study reported here.

GULP1, encoding engulfment adapter protein (Su et al., 2002), has been shown to be involved in clearance of cells undergoing apoptosis (Smits et al., 1999) during inflammation, and autoimmunity as well as embryonic development and normal tissue turnover. *GULP1* expression decreased with nicotine and levamisole treatments. Furthermore, *GULP1* showed a much more drastic decrease in expression when chronically treated with 10 μM levamisole.

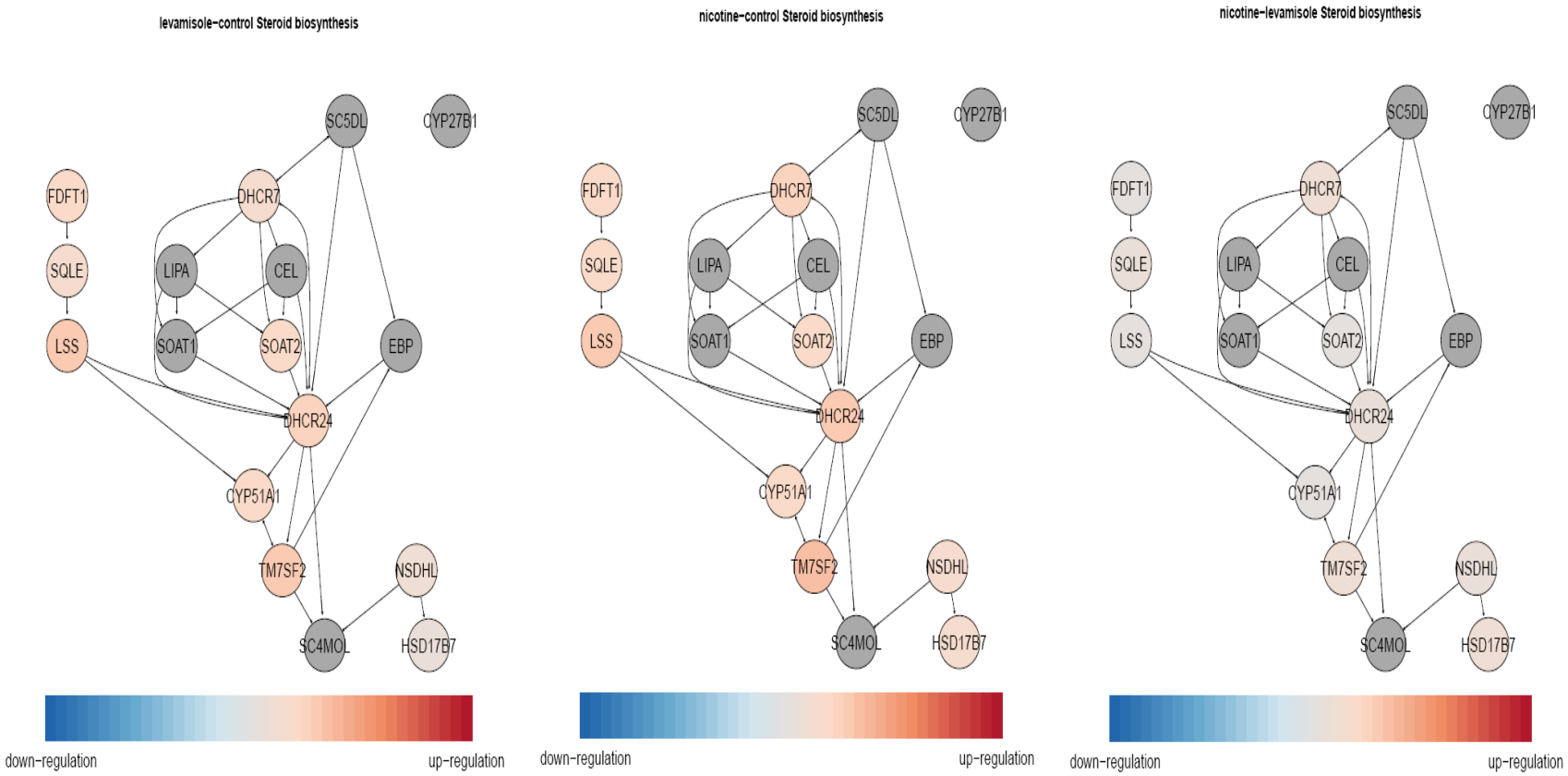

Figure 24: Steroid biosynthesis pathway in group comparisons. Generated by *Grapher* function. Differences between the expression values and fold-change ratios are powerfully represented with supportive visual graphs. *Grapher* depicts the KEGG pathway information of a set of KEGG IDs, which is the union of separate subsets obtained from separate implementations *KG_HyperG* function for different comparisons.

5. CONCLUSION AND DISCUSSION

This study uses microarray analysis methods to understand the differences and similarities between the levamisole and nicotine treatments on the SW620 colon cancer cells. The thesis aimed to A) to streamline an R-based routines based on existing and customized packages for the analysis of three group one-factor experimental design of Affymetrix arrays; B) to determine the genes commonly modulated by nicotine and levamisole; C) to determine the genes differentially modulated by nicotine and Levamisol; D) to perform functional analyses on these lists using our R-based routines; E) to confirm the expression pattern of selected genes using real-time RT-PCR of 15 day levamisole and nicotine-exposed SW620 cells.

5.1. A Flexible and Expandable Custom R Routine Designed for Automated Microarray Analysis, From QC Assessment to Functional Analysis

We have implemented a streamlined R-routine that contains multiple functions to analyze our dataset from the beginning to end, i.e., preprocessing, quality control, differential expression, functional assessment, and visualization. The programs were written in modules allowing for addition of new modules of existing and customized R routines from BioConductor and/or related resources for Affymetrix multi group one-factor analyses. Future perspectives include adding other functionalities in R environment.

BioConductor packages for statistical analysis of microarray data have already been present for more than five years, and R-based programs are frequently used for

developing custom routines. However, R packages for functional annotation and analysis of microarray data with visual support started to flourish at late 2009 and early 2010. Such existing packages provided benefits in designing and writing our R- routine. Our R-routine, although implementing similar but necessary steps many other online and standalone programs make use of in R language, it is unique because it performs all the intended analysis steps in a streamline manner, including functional analysis. Accordingly, we provide a tool which combines the necessary steps of a microarray analysis from beginning to the end, in R environment. Second, the specified R-routine has been designed to contain modules thus is highly flexible, customizable and easily expandable.

5.2. Chronic Nicotine and Levamisole Treatment Affect SW620 Colon Cancer Cells Similarly

Here in this study we show that the chronic nicotine and levamisole treatments induce similar responses in the SW620 cells. Over 1200 probes were found to be differentially expressed in the same direction with both treatments; this is striking and highly significant. It is understandable that compounds binding to similar groups of receptors could have similar overall effects, since both levamisole and nicotine are agonists of acetylcholine receptors. However, it is still interesting because L-type and N-type AChRs are known to have diverged and no cross binding is known to occur at least in nematodes (Boulin et al., 2008). Unlike the case in nematodes, there is not any levamisole-sensitive subtype of AChRs known in mammals however it has been previously shown that levamisole is a weak agonist of nAChRs; and the difference is a matter of affinity.

Our findings demonstrated that when treated chronically, levamisole and nicotine might converge in the intracellular downstream effectors of cholinergic signaling, indicated by the very similar sets of genes, molecular and cellular processes being affected by these two compounds in the same direction. On the other hand, levamisole was thought to be selective for the helminths discriminating it from the host's receptors (Rayes et al., 2004). In the light of our findings, there may not be such a selection but the timing and dosing might not be high enough to raise adverse effects on the host. Our findings suggest that treatment of 1 μM dosage chronically (for 7 days in our study) has considerable effect on the transcriptional profile of SW620 colon cancer cells and this effect is parallel to that obtained by nicotine when given in the same dosage and timing, 1 μM for seven days. Levamisole have been used as an adjuvant at high doses at which levamisole's cytotoxic effects might over dominate (Dahl et al., 2009). Our study was performed at 1uM nicotine or levamisole, mimicking the systemic blood level of nicotine in smokers (Russell et al., 1980). At this dose, the effects of nicotine and levamisole is similar; but might not be compared to the high doses used in colorectal cancer treatment.

In the chronic nicotine treatment, about 60% of the genes related to steroid biosynthesis were up-regulated along with a considerable portion of fatty acid biosynthesis, fatty acid metabolism and unsaturated fatty acid biosynthesis as well (Table 9) including pivotal enzymes of fatty acid and cholesterol synthesis, Acetyl-CoA Carboxylase (ACACA) and 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase (HMGCR). Based on these findings, it can be said that nicotine might have a considerable effect on general lipid

and steroid metabolism, at least at the cellular level. there was one very recent study implicating a role for nicotine in increased released of corticosteroids in nicotine self-administration (Yu and Sharp, 2010). Although there had been studies demonstrating a correlation between nicotine and serum cholesterol levels (Altschul et al., 1955; Ashakumary and Vijayammal, 1997; Wenzel and Beckloff, 1958), there is not any previous study established to our knowledge demonstrating a directly nicotine-induced increase in gene expression of steroid and cholesterol biosynthesis genes at the cellular level.

By no means less important, levamisole also has similar effects on lipid and steroid metabolisms. This finding is in line with the levamisole usage in steroid sensitive nephrotic syndrome, and explains how it might be possible to reduce steroids in treatment when levamisole was added (Davin and Merkus, 2005; Tanphaichitr et al., 1980). Therefore, our findings seem to shed light, at least at the cellular level, to the steroid sparing effect of levamisole in treatment of nephrotic syndrome.

5.3. Chronic nicotine and levamisole treatments increase aminoacid degradation

Interestingly, we found out degradation metabolisms of many aminoacids were up-regulated in response to chronic nicotine and levamisole treatment. These aminoacid metabolisms included valine, leucine and isoleucine degradation (~21% up-regulated), lysine degradation (~13% up-regulated), arginin and proline metabolism (~10% up-regulated), beta-alanin metabolism (~10% up-regulated) histidine metabolism (~7% up-regulated). Interestingly the most effected aminoacid metabolism was the degradation of valine, leucine and isoleucin. These three aminoacids are metabolized up until pyruvate,

an intermediary metabolite for cholesterol and lipid synthesis. Such two-sided impact might further accelerate steroid and lipid metabolisms.

These findings are more interesting in the light of a recent study on growth inhibition of constitutively active EGFR expressing glioblastoma cells by inhibiton of lipid metabolism (Guo et al., 2009). In that study, by activating AMPK, a metabolic checkpoint downstream of the LKB1 tumor suppressor and inhibitor of mTORC1 (Carracedo et al., 2008; Guo et al., 2007; Inoki et al., 2003), they show the suppression of ACACA and HMGCR to inhibit the tumor growth (Guo et al., 2009). Although, SW620 colon cancer cells are EGFR negative, lipid and steroid metabolism might still be important for membrane biosynthesis and this effect may be important in nicotine's proliferative activity much.

5.4. Anti-apoptotic or proleferation inducing genes are down-regulated in chronic nicotine and levamisole exposure

Confirmatory real-time experiments verified the decrease in anti-apoptotic gene XIAP with chronic treatment of nicotine or levamisole. Previously, nicotine, at 1 μM concentration, reported to inhibit apoptosis by up-regulating XIAP and survivin (Dasgupta et al., 2006). XIAP up-regulation was also confirmed by our group, recently, in 1 μM 48h treatment under SW620 cells under a serum starvation regime while it was not affected under normal growth conditions (Kaya, 2009). In our study we found both by analysis of microarray data and confirmed with qPCR that XIAP was downregulated with chronic exposure. Nicotine and levamisole, at 1 μM concentrations, lead to comparable decreases, yet consistent chronic exposure to 10 μM levamisole lead to a greater amount of reduction in *XIAP* expression. It is understandable that the effects of

nicotine on XIAP could differ by serum condition during the exposure to nicotine in a time dependent manner. It seems that unlike in the acute treatments, nicotine may not have an anti-apoptotic activity under the chronic treatment, however other genes involved in apoptosis should also be analyzed to summarize the effects of nicotine/levamisole effect on apoptosis in future studies. Since the change in XIAP levels are affected by serum levels as well as time of exposure to nicotine, the length of treatment periods should be carefully selected when these two compounds are studied. Previous studies have applied treatment regimes up to 14 days or longer (Konu et al., 2004). Our findings should be repeated with additional samples at day 15 to conclude that the XIAP expression is time- and dose-dependent.

The situation for *FOSL2*, a transcription factor of *AP-1* (Eferl et al., 2008), has been very similar to those seen for *XIAP*. With AP-1 interaction FOSL2 was implicated in processes like proliferation, transformation, and differentiation (Bamberger et al., 2001; Milde-Langosch et al., 2008; Nakayama et al., 2008). FOSL2 expression, was found to be up-regulated also in previous study of our group, for 48h treatment under serum starvation conditons (Kaya, 2009). In our study *FOSL2* expression was reduced. The differences between acute and chronic treatments might be differential; cell type might also be effective. Nevertheless, the differential response both between different growth conditons and between different time points is interesting. This type of flipping differential response was also seen for GULP1 and MAP1B genes. These findings higly indicate the wording for “long term” and “chronic” should be delicately used for nicotine treatment. The initial changes seen at day 7 might neuralize or might be reverted when nicotine was used for long term, i.e.> 15 days. Indeed, in the addiction

literature most experiments are performed at 14 days of exposure to nicotine at which the animals develop addiction to the drug (Konu et al., 2001). A similar time course might be appropriate for cancer cells and warrants further study.

Taken together these data suggest that nicotine may have different affects at the acute and chronic levels. The genes nicotine modulate under different conditions are rather similar; however, the direction and magnitude of changes might differ depending on the time and dose of exposures. Future studies should focus on deciphering these time and dose dependent effects on genes involved in cell proliferative and apoptotic pathways to assess the impact of smoking in cancer progression and use of levamisole in treatment of cancers. The nicotine's effects might also be driven by the microenvironment the cells are found in. Therefore, conditions such as hypoxic stress, amino acid starvation, and oxidative stress emerge as possible avenues to explore for understaing how nicotine's role in proliferation and apoptosis change.

6. FUTURE PERSPECTIVES

In this study we surveyed the effect of nicotine and levamisole on SW620 colon cancer cell line. To analyze the effect of nicotine and levamisole on gene expression we performed a microarray study with SW620 cells treated with either 1 μM nicotine or 1 μM levamisole for 7 days. For analysis of data, we designed and implemented a custom R-routine which makes extensive use of R-BioConductor Project.

This study implicates that nicotine and levamisole might have strong effect on steroid biosynthetic, cholesterol biosynthetic and aminoacid degradation metabolisms among other pathways not focused in this thesis. Confirmatory real-time PCR experiments of genes involved in steroid, cholesterol biosynthesis and valine, leucine, degradation metabolisms should be performed. Future studies should also assess nicotine's contribution in dysregulation of other metabolic and signaling pathways using confirmatory qRTPCR studies.

Our data suggest that nicotine's time dependent effects on genes involved in cell proliferation and apoptosis should be further assessed with replicated experiments performed at day 15 or later. Chronic affects of nicotine might not be apparent before within two weeks. To get clearer picture of chronic nicotine treatment and its effect on apoptosis and proliferation, apoptosis rate and viability assays should be performed along with expression analysis of a number of pro- or anti-apoptotic genes. Cell cycle analyses also would be useful to comprehend the effect of nicotine and levamisole on cell division which provide information about the silence of the cells in terms of both proliferation and apoptosis.

Our customized R-routine provides useful means for streamline automated analysis of microarray data from quality control assessments to functional annotation and analysis. First and the foremost the routine lacks an effective interface which could make it easier to use for users inexperienced in R programming language and environment. Second, the routine could be developed to include properties to perform array analyses from other species (e.g. zebrafish, mouse, rat) and include arrays produced other than Affymetrix Inc.

REFERENCES

Al-Wadei, H.A., Schuller, H.M., 2009. Nicotinic receptor-associated modulation of stimulatory and inhibitory neurotransmitters in NNK-induced adenocarcinoma of the lungs and pancreas. J Pathol 218, 437-445.

Altschul, R., Hoffer, A., Stephen, J.D., 1955. Influence of nicotinic acid on serum cholesterol in man. Arch Biochem 54, 558-559.

Arias, H.R., 1997. Topology of ligand binding sites on the nicotinic acetylcholine receptor. Brain Res Brain Res Rev 25, 133-191.

Armitage, A., Dollery, C., Houseman, T., Kohner, E., Lewis, P.J., Turner, D., 1978. Absorption of nicotine from small cigars. Clin Pharmacol Ther 23, 143-151.

Artwohl, M., Holzenbein, T., Wagner, L., Freudenthaler, A., Waldhausl, W., Baumgartner-Parzer, S.M., 2000. Levamisole induced apoptosis in cultured vascular endothelial cells. Br J Pharmacol 131, 1577-1583.

Ashakumary, L., Vijayammal, P.L., 1997. Effect of nicotine on lipoprotein metabolism in rats. Lipids 32, 311-315.

Ballivet, M., Alliod, C., Bertrand, S., Bertrand, D., 1996. Nicotinic acetylcholine receptors in the nematode Caenorhabditis elegans. J Mol Biol 258, 261-269.

Bamberger, A.M., Milde-Langosch, K., Rossing, E., Goemann, C., Loning, T., 2001. Expression pattern of the AP-1 family in endometrial cancer: correlations with cell cycle regulators. J Cancer Res Clin Oncol 127, 545-550.

Baylis, H.A., Matsuda, K., Squire, M.D., Fleming, J.T., Harvey, R.J., Darlison, M.G., Barnard, E.A., Sattelle, D.B., 1997. ACR-3, a Caenorhabditis elegans nicotinic acetylcholine receptor subunit. Molecular cloning and functional expression. Receptors Channels 5, 149-158.

Beckett, A.H., Gorrod, J.W., Jenner, P., 1970. Absorption of (-)-nicotine-l-N-oxide in man and its reduction in the gastrointestinal tract. J Pharm Pharmacol 22, 722-723.

Benowitz, N.L., 1990. Clinical pharmacology of inhaled drugs of abuse: implications in understanding nicotine dependence. NIDA Res Monogr 99, 12-29.

Benowitz, N.L., Jacob, P., 3rd, 1994. Metabolism of nicotine to cotinine studied by a dual stable isotope method. Clin Pharmacol Ther 56, 483-493.

Berg, J.Z., von Weymarn, L.B., Thompson, E.A., Wickham, K.M., Weisensel, N.A., Hatsukami, D.K., Murphy, S.E., 2010. UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption. Cancer Epidemiol Biomarkers Prev 19, 1423-1431.

Bethony, J., Brooker, S., Albonico, M., Geiger, S.M., Loukas, A., Diemert, D., Hotez, P.J., 2006. Soil-transmitted helminth infections: ascariasis, trichuriasis, and hookworm. Lancet 367, 1521-1532.

Bolstad, B.M., 2004. Low Level Analysis of High-density Oligonucleotide Array Data: Background, Normalization and Summarization. University of California, Berkeley.

Bolstad, B.M., Collin, F., Brettschneider, J., Simpson, K., Cope, L., Irizzarry, R.A., Speed, T.P., 2005. Quality Assessment of Affymetrix GeneChip Data in Bioinformatics and Computational Biology

Solutions Using R and Bioconductor., in: C.V. Gentleman R, Huber W, Irizarry R, and Dudoit S. (Ed.). Springer, New York.

Borio, G., 1995-2010. The Tobacco Timeline, Tobacco.org. Tobacco.org.

Boulin, T., Gielen, M., Richmond, J.E., Williams, D.C., Paoletti, P., Bessereau, J.L., 2008. Eight genes are required for functional reconstitution of the Caenorhabditis elegans levamisole-sensitive acetylcholine receptor. Proc Natl Acad Sci U S A 105, 18590-18595.

Brandange, S., Lindblom, L., 1979. The enzyme "aldehyde oxidase" is an iminium oxidase. Reaction with nicotine delta 1'(5') iminium ion. Biochem Biophys Res Commun 91, 991-996.

Brettschneider, J., Collin, F., Bolstad, B.M., Speed, T.P., 2007. Quality assessment for short oligonucleotide arrays. Technometrics.

Brunnemann, K.D., Hoffmann, H.D., 1974. The pH of tobacco smoke. Toxicol, 115–124.

Byrd, G.D., Chang, K.M., Greene, J.M., deBethizy, J.D., 1992. Evidence for urinary excretion of glucuronide conjugates of nicotine, cotinine, and trans-3'-hydroxycotinine in smokers. Drug Metab Dispos 20, 192-197.

Carlisle, D.L., Liu, X., Hopkins, T.M., Swick, M.C., Dhir, R., Siegfried, J.M., 2007. Nicotine activates cell-signaling pathways through muscle-type and neuronal nicotinic acetylcholine receptors in non-small cell lung cancer cells. Pulm Pharmacol Ther 20, 629-641.

Carlson, M., Falcon, S., Pages, H., Li, N., KEGG.db: A set of annotation maps for KEGG.

Carlson, M., Falcon, S., Pages, H., Li, N., R package version 2.4.1. hgu133plus2.db: Affymetrix Human Genome U133 Plus 2.0 Array annotation data (chip hgu133plus2).

Carracedo, A., Ma, L., Teruya-Feldstein, J., Rojo, F., Salmena, L., Alimonti, A., Egia, A., Sasaki, A.T., Thomas, G., Kozma, S.C., Papa, A., Nardella, C., Cantley, L.C., Baselga, J., Pandolfi, P.P., 2008. Inhibition of mTORC1 leads to MAPK pathway activation through a PI3K-dependent feedback loop in human cancer. J Clin Invest 118, 3065-3074.

Cascinu, S., Catalano, V., Piga, A., Mattioli, R., Marcellini, M., Pancotti, A., Bascioni, R., Torresi, U., Silva, R.R., Pieroni, V., Giorgi, F., Catalano, G., Cellerino, R., 2003. The role of levamisole in the adjuvant treatment of stage III colon cancer patients: a randomized trial of 5-fluorouracil and levamisole versus 5-fluorouracil alone. Cancer Invest 21, 701-707.

Cashman, J.R., Park, S.B., Yang, Z.C., Wrighton, S.A., Jacob, P., 3rd, Benowitz, N.L., 1992. Metabolism of nicotine by human liver microsomes: stereoselective formation of trans-nicotine N'-oxide. Chem Res Toxicol 5, 639-646.

Cecchin, E., Innocenti, F., D'Andrea, M., Corona, G., De Mattia, E., Biason, P., Buonadonna, A., Toffoli, G., 2009. Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. J Clin Oncol 27, 2457-2465.

Changeux, J.P., Edelstein, S.J., 1998. Allosteric receptors after 30 years. Neuron 21, 959-980.

Culetto, E., Baylis, H.A., Richmond, J.E., Jones, A.K., Fleming, J.T., Squire, M.D., Lewis, J.A., Sattelle, D.B., 2004. The Caenorhabditis elegans unc-63 gene encodes a levamisole-sensitive nicotinic acetylcholine receptor alpha subunit. J Biol Chem 279, 42476-42483.

Culhane, A., made4: Multivariate analysis of microarray data using ADE4. R package version 1.22.0.

Dahl, O., Fluge, O., Carlsen, E., Wiig, J.N., Myrvold, H.E., Vonen, B., Podhorny, N., Bjerkeset, O., Eide, T.J., Halvorsen, T.B., Tveit, K.M., 2009. Final results of a randomised phase III study on adjuvant chemotherapy with 5 FU and levamisol in colon and rectum cancer stage II and III by the Norwegian Gastrointestinal Cancer Group. Acta Oncol 48, 368-376.

Dajani, R.M., Gorrod, J.W., Beckett, A.H., 1975. Reduction in vivo of (minus)-nicotine-1'-N-oxide by germ-free and conventional rats. Biochem Pharmacol 24, 648-650.

Dajas-Bailador, F., Wonnacott, S., 2004. Nicotinic acetylcholine receptors and the regulation of neuronal signalling. Trends Pharmacol Sci 25, 317-324.

Dasgupta, P., Kinkade, R., Joshi, B., Decook, C., Haura, E., Chellappan, S., 2006. Nicotine inhibits apoptosis induced by chemotherapeutic drugs by up-regulating XIAP and survivin. Proc Natl Acad Sci U S A 103, 6332-6337.

Davin, J.C., Merkus, M.P., 2005. Levamisole in steroid-sensitive nephrotic syndrome of childhood: the lost paradise? Pediatr Nephrol 20, 10-14.

Dayal, U., Dayal, A.K., Shastry, J.C., Raghupathy, P., 1994. Use of levamisole in maintaining remission in steroid-sensitive nephrotic syndrome in children. Nephron 66, 408-412.

De Placido, S., Lopez, M., Carlomagno, C., Paoletti, G., Palazzo, S., Manzione, L., Iannace, C., Ianniello, G.P., De Vita, F., Ficorella, C., Farris, A., Pistillucci, G., Gemini, M., Cortesi, E., Adamo, V., Gebbia, N., Palmeri, S., Gallo, C., Perrone, F., Persico, G., Bianco, A.R., 2005. Modulation of 5-fluorouracil as adjuvant systemic chemotherapy in colorectal cancer: the IGCS-COL multicentre, randomised, phase III study. Br J Cancer 93, 896-904.

Derby, K.S., Cuthrell, K., Caberto, C., Carmella, S., Murphy, S.E., Hecht, S.S., Le Marchand, L., 2009. Exposure to the carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in smokers from 3 populations with different risks of lung cancer. Int J Cancer 125, 2418-2424.

Doll, R., Hill, A.B., 1950. Smoking and carcinoma of the lung; preliminary report. Br Med J 2, 739-748.

Drell, T.L.t., Joseph, J., Lang, K., Niggemann, B., Zaenker, K.S., Entschladen, F., 2003. Effects of neurotransmitters on the chemokinesis and chemotaxis of MDA-MB-468 human breast carcinoma cells. Breast Cancer Res Treat 80, 63-70.

Dwoskin, L.P., Leibee, L.L., Jewell, A.L., Fang, Z.X., Crooks, P.A., 1992. Inhibition of [3H]dopamine uptake into rat striatal slices by quaternary N-methylated nicotine metabolites. Life Sci 50, PL233-237.

Eferl, R., Hasselblatt, P., Rath, M., Popper, H., Zenz, R., Komnenovic, V., Idarraga, M.H., Kenner, L., Wagner, E.F., 2008. Development of pulmonary fibrosis through a pathway involving the transcription factor Fra-2/AP-1. Proc Natl Acad Sci U S A 105, 10525-10530.

Eimer, S., Gottschalk, A., Hengartner, M., Horvitz, H.R., Richmond, J., Schafer, W.R., Bessereau, J.L., 2007. Regulation of nicotinic receptor trafficking by the transmembrane Golgi protein UNC-50. EMBO J 26, 4313-4323.

Essenberg, J.M., Horowitz, M., Gaffney, E., 1955. The incidence of lung tumors in albino mice exposed to the smoke from cigarettes low in nicotine content. West J Surg Obstet Gynecol 63, 265-267.

Feng, G., Du, P., Kibbe, W., Lin, S., 2001. GeneAnswers: Integrated Interpretation of Genes.

Finnegan, J.K., Larson, P.S., Haag, H.B., 1945. The Role of Nicotine in the Cigarette Habit. Science 102, 94-96.

Fleming, J.T., Squire, M.D., Barnes, T.M., Tornoe, C., Matsuda, K., Ahnn, J., Fire, A., Sulston, J.E., Barnard, E.A., Sattelle, D.B., Lewis, J.A., 1997. Caenorhabditis elegans levamisole resistance genes lev-1, unc-29, and unc-38 encode functional nicotinic acetylcholine receptor subunits. J Neurosci 17, 5843-5857.

Fowler, R.T., 1954. A Redetermination of ionization constants of nicotine. J Appl Chem, 449-452.

Fu, X.W., Lindstrom, J., Spindel, E.R., 2009. Nicotine activates and up-regulates nicotinic acetylcholine receptors in bronchial epithelial cells. Am J Respir Cell Mol Biol 41, 93-99.

Gaimarri, A., Moretti, M., Riganti, L., Zanardi, A., Clementi, F., Gotti, C., 2007. Regulation of neuronal nicotinic receptor traffic and expression. Brain Res Rev 55, 134-143.

Garchon, H.J., Djabiri, F., Viard, J.P., Gajdos, P., Bach, J.F., 1994. Involvement of human muscle acetylcholine receptor alpha-subunit gene (CHRNA) in susceptibility to myasthenia gravis. Proc Natl Acad Sci U S A 91, 4668-4672.

Gautier, L., Cope, L., Bolstad, B.M., Irizarry, R.A., 2004. *affy* -- analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20, 307-315.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y., Zhang, J., 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5, R80.

Gentry, J., Long, L., Gentleman, R., Falcon, S., Hahne, F., Sarkar, D., Hansen, K., Rgraphviz: Provides plotting capabilities for R graph objects. R package version 1.26.0.

Gori, G.B., Benowitz, N.L., Lynch, C.J., 1986. Mouth versus deep airways absorption of nicotine in cigarette smokers. Pharmacol Biochem Behav 25, 1181-1184.

Gotti, C., Clementi, F., 2004. Neuronal nicotinic receptors: from structure to pathology. Prog Neurobiol 74, 363-396.

Green, W.N., Millar, N.S., 1995. Ion-channel assembly. Trends Neurosci 18, 280-287.

Guo, D., Chien, S., Shyy, J.Y., 2007. Regulation of endothelial cell cycle by laminar versus oscillatory flow: distinct modes of interactions of AMP-activated protein kinase and Akt pathways. Circ Res 100, 564-571.

Guo, D., Hildebrandt, I.J., Prins, R.M., Soto, H., Mazzotta, M.M., Dang, J., Czernin, J., Shyy, J.Y., Watson, A.D., Phelps, M., Radu, C.G., Cloughesy, T.F., Mischel, P.S., 2009. The AMPK agonist AICAR inhibits the growth of EGFRvIII-expressing glioblastomas by inhibiting lipogenesis. Proc Natl Acad Sci U S A 106, 12932-12937.

Halevi, S., McKay, J., Palfreyman, M., Yassin, L., Eshel, M., Jorgensen, E., Treinin, M., 2002. The C. elegans ric-3 gene is required for maturation of nicotinic acetylcholine receptors. EMBO J 21, 1012-1020.

Halevi, S., Yassin, L., Eshel, M., Sala, F., Sala, S., Criado, M., Treinin, M., 2003. Conservation within the RIC-3 gene family. Effectors of mammalian nicotinic acetylcholine receptor expression. J Biol Chem 278, 34411-34417.

Hatsukami, D.K., Lemmonds, C., Zhang, Y., Murphy, S.E., Le, C., Carmella, S.G., Hecht, S.S., 2004. Evaluation of carcinogen exposure in people who used "reduced exposure" tobacco products. J Natl Cancer Inst 96, 844-852.

Hecht, S.S., 2002. Human urinary carcinogen metabolites: biomarkers for investigating tobacco and cancer. Carcinogenesis 23, 907-922.

Hecht, S.S., Hochalter, J.B., Villalta, P.W., Murphy, S.E., 2000. 2'-Hydroxylation of nicotine by cytochrome P450 2A6 and human liver microsomes: formation of a lung carcinogen precursor. Proc Natl Acad Sci U S A 97, 12493-12497.

Heeschen, C., Weis, M., Aicher, A., Dimmeler, S., Cooke, J.P., 2002. A novel angiogenic pathway mediated by non-neuronal nicotinic acetylcholine receptors. J Clin Invest 110, 527-536.

HelminthInfections, 2010. Helminth Infections, http://helminth-infections.com/. Copyright © Helminth Infections.

Hoffmann, D., Hoffmann, I., El-Bayoumy, K., 2001. The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder. Chem Res Toxicol 14, 767-790.

Holcombe, R.F., Li, A., Stewart, R.M., 1998. Levamisole and interleukin-2 for advanced malignancy. Biotherapy 11, 255-258.

http://bioinfo.vanderbilt.edu/webgestalt/, WebGestalt: Web-based Gene Set Analysis Toolkit.

http://bioinforx.com/, BioInfoRx LIMS Laboratory Information Management Systems and Automated Data Analysis.

http://david.abcc.ncifcrf.gov/, DAVID Functional Annotation Bioinformatics Microarray Analysis.

http://www.biorag.org/, Bio Resource for Array Genes.

http://www.broadinstitute.org/gsea/, Gene Set Enrichment Analysis.

http://www.broadinstitute.org/gsea/msigdb/, GSEA | MSigDB.

http://www.ingenuity.com/, Ingenuity Pathway Analysis Software-Complete Pathways Database.

https://www.emaas.org/EMAAS/, EMAAS - Extensible MicroArray Analysis System.

Hukkanen, J., Jacob, P., 3rd, Benowitz, N.L., 2005. Metabolism and disposition kinetics of nicotine. Pharmacol Rev 57, 79-115.

Inoki, K., Zhu, T., Guan, K.L., 2003. TSC2 mediates cellular energy response to control cell growth and survival. Cell 115, 577-590.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., Speed, T.P., 2003. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31, e15.

Jensen, A.A., Frolund, B., Liljefors, T., Krogsgaard-Larsen, P., 2005. Neuronal nicotinic acetylcholine receptors: structural revelations, target identifications, and therapeutic inspirations. J Med Chem 48, 4705-4745.

Jones, A.K., Sattelle, D.B., 2004. Functional genomics of the nicotinic acetylcholine receptor gene family of the nematode, Caenorhabditis elegans. Bioessays 26, 39-49.

Joseph, J., Niggemann, B., Zaenker, K.S., Entschladen, F., 2002. The neurotransmitter gamma-aminobutyric acid is an inhibitory regulator for the migration of SW 480 colon carcinoma cells. Cancer Res 62, 6467-6469.

Kaivosaari, S., Toivonen, P., Hesse, L.M., Koskinen, M., Court, M.H., Finel, M., 2007. Nicotine glucuronidation and the human UDP-glucuronosyltransferase UGT2B10. Mol Pharmacol 72, 761-768.

Kawai, H., Berg, D.K., 2001. Nicotinic acetylcholine receptors containing alpha 7 subunits on rat cortical neurons do not undergo long-lasting inactivation even when up-regulated by chronic nicotine exposure. J Neurochem 78, 1367-1378.

Kaya, O., 2009. Investigations of the effects of nicotine on the expression profile of SW620 colon adenocarcinoma cells using a functional genomics approach, Dept. of Molecular Biology and Genetics. Bilkent University, Ankara, 96.

Kim, J.A., Kang, Y.S., Lee, S.H., Lee, E.H., Lee, Y.S., 2001. Role of pertussis toxin-sensitive G-proteins in intracellular Ca2+ release and apoptosis induced by inhibiting cystic fibrosis transmembrane conductance regulator (CFTR) Cl- channels in HepG2 human hepatoblastoma cells. J Cell Biochem 81, 93-101.

Kim, J.A., Kang, Y.S., Lee, S.H., Lee, E.H., Yoo, B.H., Lee, Y.S., 1999. Glibenclamide induces apoptosis through inhibition of cystic fibrosis transmembrane conductance regulator (CFTR) Cl(-) channels and intracellular Ca(2+) release in HepG2 human hepatoblastoma cells. Biochem Biophys Res Commun 261, 682-688.

Kittler, J.T., Moss, S.J., 2001. Neurotransmitter receptor trafficking and the regulation of synaptic strength. Traffic 2, 437-448.

Konu, O., Kane, J.K., Barrett, T., Vawter, M.P., Chang, R., Ma, J.Z., Donovan, D.M., Sharp, B., Becker, K.G., Li, M.D., 2001. Region-specific transcriptional response to chronic nicotine in rat brain. Brain Res 909, 194-203.

Konu, O., Xu, X., Ma, J.Z., Kane, J., Wang, J., Shi, S.J., Li, M.D., 2004. Application of a customized pathway-focused microarray for gene expression profiling of cellular homeostasis upon exposure to nicotine in PC12 cells. Brain Res Mol Brain Res 121, 102-113.

Kuehl, G.E., Murphy, S.E., 2003. N-glucuronidation of nicotine and cotinine by human liver microsomes and heterologously expressed UDP-glucuronosyltransferases. Drug Metab Dispos 31, 1361-1368.

Kuryatov, A., Luo, J., Cooper, J., Lindstrom, J., 2005. Nicotine acts as a pharmacological chaperone to up-regulate human alpha4beta2 acetylcholine receptors. Mol Pharmacol 68, 1839-1851.

Laurie, J.A., Moertel, C.G., Fleming, T.R., Wieand, H.S., Leigh, J.E., Rubin, J., McCormack, G.W., Gerstner, J.B., Krook, J.E., Malliard, J., et al., 1989. Surgical adjuvant therapy of large-bowel carcinoma: an evaluation of levamisole and the combination of levamisole and fluorouracil. The North Central Cancer Treatment Group and the Mayo Clinic. J Clin Oncol 7, 1447-1456.

Leibovici, J., Itzhaki, O., Kaptzan, T., Skutelsky, E., Sinai, J., Michowitz, M., Asfur, R., Siegal, A., Huszar, M., Schiby, G., 2009. Designing ageing conditions in tumour microenvironment-a new possible modality for cancer treatment. Mech Ageing Dev 130, 76-85.

Lena, C., Changeux, J.P., 1998. Allosteric nicotinic receptors, human pathologies. J Physiol Paris 92, 63-74.

Lessov-Schlaggar, C.N., Benowitz, N.L., Jacob, P., Swan, G.E., 2009. Genetic influences on individual differences in nicotine glucuronidation. Twin Res Hum Genet 12, 507-513.

Lester, H.A., Xiao, C., Srinivasan, R., Son, C.D., Miwa, J., Pantoja, R., Banghart, M.R., Dougherty, D.A., Goate, A.M., Wang, J.C., 2009. Nicotine is a selective pharmacological chaperone of acetylcholine receptor number and stoichiometry. Implications for drug discovery. AAPS J 11, 167-177.

Levandoski, M.M., Piket, B., Chang, J., 2003. The anthelmintic levamisole is an allosteric modulator of human neuronal nicotinic acetylcholine receptors. Eur J Pharmacol 471, 9-20.

Levin, M.L., Goldstein, H., Gerhardt, P.R., 1950. Cancer and tobacco smoking; a preliminary report. J Am Med Assoc 143, 336-338.

Lewis, J.A., Elmer, J.S., Skimming, J., McLafferty, S., Fleming, J., McGee, T., 1987a. Cholinergic receptor mutants of the nematode Caenorhabditis elegans. J Neurosci 7, 3059-3071.

Lewis, J.A., Fleming, J.T., McLafferty, S., Murphy, H., Wu, C., 1987b. The levamisole receptor, a cholinergic receptor of the nematode Caenorhabditis elegans. Mol Pharmacol 31, 185-193.

Lewis, J.A., Wu, C.H., Berg, H., Levine, J.H., 1980. The genetics of levamisole resistance in the nematode Caenorhabditis elegans. Genetics 95, 905-928.

Lindstrom, J.M., 2000. Acetylcholine receptors and myasthenia. Muscle Nerve 23, 453-477.

Luetje, C.W., Patrick, J., 1991. Both alpha- and beta-subunits contribute to the agonist sensitivity of neuronal nicotinic acetylcholine receptors. J Neurosci 11, 837-845.

McKennis, H., Jr., Turnbull, L.B., Bowman, E.R., 1963. N-Methylation of nicotine ad cotinine in vivo. J Biol Chem 238, 719-723.

Milde-Langosch, K., Janke, S., Wagner, I., Schroder, C., Streichert, T., Bamberger, A.M., Janicke, F., Loning, T., 2008. Role of Fra-2 in breast cancer: influence on tumor cell invasion and motility. Breast Cancer Res Treat 107, 337-347.

Moertel, C.G., Fleming, T.R., Macdonald, J.S., Haller, D.G., Laurie, J.A., Goodman, P.J., Ungerleider, J.S., Emerson, W.A., Tormey, D.C., Glick, J.H., et al., 1990. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. N Engl J Med 322, 352-358.

Mongan, N.P., Baylis, H.A., Adcock, C., Smith, G.R., Sansom, M.S., Sattelle, D.B., 1998. An extensive and diverse gene family of nicotinic acetylcholine receptor alpha subunits in Caenorhabditis elegans. Receptors Channels 6, 213-228.

Mongan, N.P., Jones, A.K., Smith, G.R., Sansom, M.S., Sattelle, D.B., 2002. Novel alpha7-like nicotinic acetylcholine receptor subunits in the nematode Caenorhabditis elegans. Protein Sci 11, 1162-1171.

Murphy, P.J., 1973. Enzymatic oxidation of nicotine to nicotine 1'(5') iminium ion. A newly discovered intermediate in the metabolism of nicotine. J Biol Chem 248, 2796-2800.

Nakajima, M., Yamamoto, T., Nunoya, K., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N., Kamataki, T., Kuroiwa, Y., 1996. Characterization of CYP2A6 involved in 3'-hydroxylation of cotinine in human liver microsomes. J Pharmacol Exp Ther 277, 1010-1015.

Nakayama, T., Hieshima, K., Arao, T., Jin, Z., Nagakubo, D., Shirakawa, A.K., Yamada, Y., Fujii, M., Oiso, N., Kawada, A., Nishio, K., Yoshie, O., 2008. Aberrant expression of Fra-2 promotes CCR4 expression and cell proliferation in adult T-cell leukemia. Oncogene 27, 3221-3232.

Neurath, G., Orth, D., Pein, F., 1991. Detection of nornicotine in human urine after infusion of nicotine, in: F. Adlkofer, K. Thurau (Eds.), Effects of Nicotine on Biological Systems, Birkhauser Verlag, Basel, 45-49.

NIH_Consensus_Conference, 1990. Adjuvant therapy for patients with colon and rectal cancer. JAMA 264, 1444-1450.

O'Connell, M.J., Laurie, J.A., Kahn, M., Fitzgibbons, R.J., Jr., Erlichman, C., Shepherd, L., Moertel, C.G., Kocha, W.I., Pazdur, R., Wieand, H.S., Rubin, J., Vukov, A.M., Donohue, J.H., Krook, J.E., Figueredo, A., 1998. Prospectively randomized trial of postoperative adjuvant chemotherapy in patients with high-risk colon cancer. J Clin Oncol 16, 295-300.

Pakkanen, J.S., Stenfors, J., Jokitalo, E., Tuominen, R.K., 2006. Effect of chronic nicotine treatment on localization of neuronal nicotinic acetylcholine receptors at cellular level. Synapse 59, 383-393.

Pankow, J.F., 2001. A consideration of the role of gas/particle partitioning in the deposition of nicotine and other tobacco smoke compounds in the respiratory tract. Chem Res Toxicol 14, 1465-1481.

Pankow, J.F., Tavakoli, A.D., Luo, W., Isabelle, L.M., 2003. Percent free base nicotine in the tobacco smoke particulate matter of selected commercial and reference cigarettes. Chem Res Toxicol 16, 1014-1018.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29, e45.

PMEP, 1985. nicotine(Black Leaf 40) Chemical Profile 4/85, nicotine(Black Leaf 40) Chemical Profile 4/85, 4 23 ed. Cornell University.

Quiram, P.A., Ohno, K., Milone, M., Patterson, M.C., Pruitt, N.J., Brengman, J.M., Sine, S.M., Engel, A.G., 1999. Mutation causing congenital myasthenia reveals acetylcholine receptor beta/delta subunit interaction essential for assembly. J Clin Invest 104, 1403-1410.

R_Development_Core_Team, 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rainer, J., Sanchez-Cabo, F., Stocker, G., Sturn, A., Trajanoski, Z., 2006. CARMAweb: comprehensive R- and bioconductor-based web service for microarray data analysis. Nucleic Acids Res 34, W498-503.

Rashid, H.U., Ahmed, S., Fatima, N., Khanam, A., 1996. Levamisole in the treatment of steroid dependent or frequent relapsing nephritic syndrome in children. Bangladesh Ren J, 6-8.

Rayes, D., De Rosa, M.J., Bartos, M., Bouzat, C., 2004. Molecular basis of the differential sensitivity of nematode and mammalian muscle to the anthelmintic agent levamisole. J Biol Chem 279, 36372-36381.

Rayes, D., Flamini, M., Hernando, G., Bouzat, C., 2007. Activation of single nicotinic receptor channels from Caenorhabditis elegans muscle. Mol Pharmacol 71, 1407-1415.

Ritter, J.K., 2000. Roles of glucuronidation and UDP-glucuronosyltransferases in xenobiotic bioactivation reactions. Chem Biol Interact 129, 171-193.

Roberts, G., IPCS INCHEM. IPCS INCHEM, International Programme on Chemical Safety. Chemical Safety Information from Intergovernmental Organizations, in: G. Roberts (Ed.), http://www.inchem.org/. (IPCS), ® Copyright International Programme on Chemical Safety; (CCOHS), Canadian Centre for Occupational Health and Safety;.

Rozen, S., Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132, 365-386.

Russell, M.A., Jarvis, M., Iyer, R., Feyerabend, C., 1980. Relation of nicotine yield of cigarettes to blood nicotine concentrations in smokers. Br Med J 280, 972-976.

Schoenfeld, T.A., McKerracher, L., Obar, R., Vallee, R.B., 1989. MAP 1A and MAP 1B are structurally related microtubule associated proteins with distinct developmental patterns in the CNS. J Neurosci 9, 1712-1730.

Schuller, H.M., 2009. Is cancer triggered by altered signalling of nicotinic acetylcholine receptors? Nat Rev Cancer 9, 195-205.

Schuller, H.M., Al-Wadei, H.A., Majidi, M., 2008a. GABA B receptor is a novel drug target for pancreatic cancer. Cancer 112, 767-778.

Schuller, H.M., Al-Wadei, H.A., Majidi, M., 2008b. Gamma-aminobutyric acid, a potential tumor suppressor for small airway-derived lung adenocarcinoma. Carcinogenesis 29, 1979-1985.

Scott, F.L., Denault, J.B., Riedl, S.J., Shin, H., Renatus, M., Salvesen, G.S., 2005. XIAP inhibits caspase-3 and -7 using two binding sites: evolutionarily conserved mechanism of IAPs. EMBO J 24, 645-655.

Seaton, M.J., Vesell, E.S., Luo, H., Hawes, E.M., 1993. Identification of radiolabeled metabolites of nicotine in rat bile. Synthesis of S-(-)-nicotine N-glucuronide and direct separation of nicotine-derived conjugates using high-performance liquid chromatography. J Chromatogr 621, 49-53.

Sensabaugh, A.J., Cundiff, R.H., 1967. A new technique for determining the pH of whole tobacco smoke. Tobacco Sci, 28–33.

Smits, E., Van Criekinge, W., Plaetinck, G., Bogaert, T., 1999. The human homologue of Caenorhabditis elegans CED-6 specifically promotes phagocytosis of apoptotic cells. Curr Biol 9, 1351-1354.

Smyth, G.K., 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3, Article3.

Smyth, G.K., 2005. Limma: linear models for microarray data, in: W.H. R. Gentleman and V. Carey and S. Dudoit and R. Irizarry (Ed.), Bioinformatics and Computational Biology Solutions using R and Bioconductor. Springer, New York, 397-420.

Squire, M.D., Tornoe, C., Baylis, H.A., Fleming, J.T., Barnard, E.A., Sattelle, D.B., 1995. Molecular cloning and functional co-expression of a Caenorhabditis elegans nicotinic acetylcholine receptor subunit (acr-2). Receptors Channels 3, 107-115.

Su, H.P., Nakada-Tsukui, K., Tosello-Trampont, A.C., Li, Y., Bu, G., Henson, P.M., Ravichandran, K.S., 2002. Interaction of CED-6/GULP, an adapter protein involved in engulfment of apoptotic cells with CED-1 and CD91/low density lipoprotein receptor-related protein (LRP). J Biol Chem 277, 11772-11779.

Sun, B., Sterling, C.R., Tank, A.W., 2003. Chronic nicotine treatment leads to sustained stimulation of tyrosine hydroxylase gene transcription rate in rat adrenal medulla. J Pharmacol Exp Ther 304, 575-588.

Tanphaichitr, P., Tanphaichitr, D., Sureeratanan, J., Chatasingh, S., 1980. Treatment of nephrotic syndrome with levamisole. J Pediatr 96, 490-493.

Treinin, M., Chalfie, M., 1995. A mutated acetylcholine receptor subunit causes neuronal degeneration in C. elegans. Neuron 14, 871-877.

Treinin, M., Gillo, B., Liebman, L., Chalfie, M., 1998. Two functionally dependent acetylcholine subunits are encoded in a single Caenorhabditis elegans operon. Proc Natl Acad Sci U S A 95, 15492-15495.

Upadhyaya, P., Zimmerman, C.L., Hecht, S.S., 2002. Metabolism and pharmacokinetics of N'-nitrosonornicotine in the patas monkey. Drug Metab Dispos 30, 1115-1122.

van der Bol, J.M., Visser, T.J., Loos, W.J., de Jong, F.A., Wiemer, E.A., van Aken, M.O., Planting, A.S., Schellens, J.H., Verweij, J., Mathijssen, R.H., 2010. Effects of methimazole on the elimination of irinotecan. Cancer Chemother Pharmacol.

Wang, Y., McClelland, M., Xia, X.Q., 2009. Analyzing microarray data using WebArray. Cold Spring Harb Protoc 2009, pdb prot5260.

Wenzel, D.G., Beckloff, G.L., 1958. The effect of nicotine on experimental hypercholesterolemia in the rabbit. J Am Pharm Assoc Am Pharm Assoc (Baltim) 47, 338-343.

Wong, H.P., Yu, L., Lam, E.K., Tai, E.K., Wu, W.K., Cho, C.H., 2007. Nicotine promotes cell proliferation via alpha7-nicotinic acetylcholine receptor and catecholamine-synthesizing enzymes-mediated pathway in human colon adenocarcinoma HT-29 cells. Toxicol Appl Pharmacol 221, 261-267.

WormBase, 2010. WormBase Web Site.

www.babelomics.org, Babelomics 4.

www.bioconductor.org, BioConductor.

Wynder, E.L., Graham, E.A., 1950. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma; a study of 684 proved cases. J Am Med Assoc 143, 329-336.

Xia, X., McClelland, M., Wang, Y., 2005. WebArray: an online platform for microarray data analysis. BMC Bioinformatics 6, 306.

Xia, X.Q., McClelland, M., Porwollik, S., Song, W., Cong, X., Wang, Y., 2009. WebArrayDB: cross-platform microarray data analysis and public data repository. Bioinformatics 25, 2425-2429.

Yakoub, A., Gustafsson, L., Ericsson, O., Hellgren, U., 1995. Handbook of Drugs for Tropical Parasitic Infections, Bristol.

Yamanaka, H., Nakajima, M., Fukami, T., Sakai, H., Nakamura, A., Katoh, M., Takamiya, M., Aoki, Y., Yokoi, T., 2005a. CYP2A6 AND CYP2B6 are involved in nornicotine formation from nicotine in humans: interindividual differences in these contributions. Drug Metab Dispos 33, 1811-1818.

Yamanaka, H., Nakajima, M., Katoh, M., Kanoh, A., Tamura, O., Ishibashi, H., Yokoi, T., 2005b. Trans-3'-hydroxycotinine O- and N-glucuronidations in human liver microsomes. Drug Metab Dispos 33, 23-30.

Yu, G., Sharp, B.M., 2010. Nicotine self-administration diminishes stress-induced norepinephrine secretion but augments adrenergic-responsiveness in the hypothalamic paraventricular nucleus and enhances adrenocorticotropic hormone and corticosterone release. J Neurochem 112, 1327-1337.

Zhang, J.D., Wiemann, S., 2009. KEGGgraph: KEGGgraph: A graph approach to KEGG PATHWAY in R and Bioconductor.

Zhang, W., Du, X., Zhao, G., Jin, H., Kang, Y., Xiao, C., Liu, M., Wang, B., 2009. Levamisole is a potential facilitator for the activation of Th1 responses of the subunit HBV vaccination. Vaccine 27, 4938-4946.

APPENDIX

Table 11: Full *KEGGscript* results.

GROUPS	Levamisole- Control		Nicotine-Control		Nicotine-Levamisole	
PATHWAY NAME	**DOWN %**	**UP %**	**DOWN %**	**UP %**	**DOWN %**	**UP %**
Steroid biosynthesis	-	47.06	-	58.82	-	5.88
Terpenoid backbone biosynthesis	-	26.67	-	26.67	-	-
Limonene and pinene degradation	-	25.00	-	25.00	-	-
Lysine biosynthesis	-	-	-	25.00	-	-
Propanoate metabolism	3.03	24.24	-	24.24	-	-
Valine, leucine and isoleucine degradation	-	22.73	-	20.45	-	2.27
Pentose phosphate pathway	-	14.81	-	18.52	-	-
Glycerolipid metabolism	4.08	14.29	4.08	18.37	-	-
Glutathione metabolism	-	1-	-	18.00	-	-
Pyruvate metabolism	2.50	15.00	-	17.50	-	-
Fatty acid biosynthesis	-	16.67	-	16.67	-	-
Citrate cycle (TCA cycle)	-	9.68	-	16.13	-	-
Fatty acid metabolism	-	16.67	-	14.29	-	-
Ubiquinone and other terpenoid-quinone biosynthesis	-	14.29	-	14.29	-	-
Lysine degradation	2.17	10.87	2.17	13.04	-	-
Nicotinate and nicotinamide metabolism	4.17	8.33	4.17	12.50	4.17	-
Tryptophan metabolism	2.50	1-	2.50	12.50	-	-
Thiamine metabolism	-	12.50	12.50	12.50	-	-
Fatty acid elongation in mitochondria	-	-	-	12.50	-	-
Galactose metabolism	3.85	7.69	3.85	11.54	3.85	-
Butanoate metabolism	-	14.29	-	11.43	-	-
Synthesis and degradation of ketone bodies	-	11.11	-	11.11	-	-
Glycolysis / Gluconeogenesis	1.56	10.94	-	10.94	-	-
Biosynthesis of unsaturated fatty acids	4.76	14.29	-	9.52	-	-
Arginine and proline metabolism	1.85	9.26	3.70	9.26	-	-
beta-Alanine metabolism	4.55	22.73	4.55	9.09	-	-
Valine, leucine and isoleucine biosynthesis	9.09	9.09	-	9.09	-	-
Glycerophospholipid metabolism	1.27	7.59	2.53	8.86	-	-
Fructose and mannose metabolism	-	8.82	-	8.82	-	-
Inositol phosphate metabolism	5.26	5.26	8.77	8.77	-	1.75
Nitrogen metabolism	-	8.70	4.35	8.70	-	-
Ether lipid metabolism	-	8.57	2.86	8.57	-	-

Metabolic pathways	1.96	7.03	2.67	8.19	0.18	0.53
Ascorbate and aldarate metabolism	4.00	8.00	4.00	8.00	-	-
Phosphatidylinositol signaling system	5.13	6.41	7.69	7.69	-	2.56
Sulfur metabolism	-	-	-	7.69	-	-
Steroid hormone biosynthesis	3.64	-	3.64	7.27	-	-
Pyrimidine metabolism	3.06	5.10	2.04	7.14	-	-
Histidine metabolism	-	10.34	-	6.90	-	-
Riboflavin metabolism	6.25	6.25	-	6.25	6.25	-
Primary bile acid biosynthesis	6.25	6.25	6.25	6.25	-	-
Pantothenate and CoA biosynthesis	-	6.25	-	6.25	6.25	-
Glyoxylate and dicarboxylate metabolism	-	6.25	-	6.25	-	-
Other glycan degradation	-	-	-	6.25	-	-
Cell cycle	3.48	5.22	1.74	6.09	0.87	2.61
One carbon pool by folate	5.88	17.65	-	5.88	-	5.88
p53 signaling pathway	12.86	8.57	7.14	5.71	-	4.29
Metabolism of xenobiotics by cytochrome P450	2.86	2.86	2.86	5.71	-	-
Drug metabolism - cytochrome P450	-	2.78	-	5.56	-	-
Pancreatic cancer	1.37	4.11	2.74	5.48	-	2.74
Wnt signaling pathway	8.11	4.05	4.05	5.41	0.68	2.70
Tyrosine metabolism	2.44	2.44	2.44	4.88	-	-
Colorectal cancer	4.76	3.57	3.57	4.76	1.19	4.76
Apoptosis	5.88	4.71	1.18	4.71	-	5.88
ErbB signaling pathway	1.15	5.75	4.60	4.60	1.15	3.45
Glycosaminoglycan biosynthesis - chondroitin sulfate	9.09	4.55	13.64	4.55	-	-
Oxidative phosphorylation	-	2.99	0.75	4.48	-	0.75
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	-	4.00	4.00	4.00	-	-
Purine metabolism	3.14	2.52	1.89	3.77	0.63	0.63
Pentose and glucuronate interconversions	3.57	7.14	3.57	3.57	-	-
Glycosphingolipid biosynthesis - lacto and neolacto series	3.57	3.57	3.57	3.57	-	-
Arachidonic acid metabolism	-	1.72	-	3.45	-	-
O-Glycan biosynthesis	-	1-	3.33	3.33	-	-
Glycine, serine and threonine metabolism	3.23	-	3.23	3.23	-	-
Alanine, aspartate and glutamate metabolism	6.25	3.13	6.25	3.13	-	-
Retinol metabolism	-	1.56	1.56	3.13	-	-
VEGF signaling pathway	4.11	4.11	2.74	2.74	-	2.74
MAPK signaling pathway	3.40	4.15	3.02	2.64	0.75	1.13
Sphingolipid metabolism	-	2.50	2.50	2.50	-	-
Aminoacyl-tRNA biosynthesis	4.88	-	2.44	2.44	-	2.44

Porphyrin and chlorophyll metabolism	-	2.38	2.38	2.38	-	-
Amino sugar and nucleotide sugar metabolism	6.82	2.27	6.82	2.27	-	2.27
N-Glycan biosynthesis	4.17	2.08	6.25	2.08	-	-
Drug metabolism - other enzymes	3.92	-	1.96	1.96	-	-
Starch and sucrose metabolism	1.92	-	1.92	1.92	3.85	-
Jak-STAT signaling pathway	3.87	1.94	4.52	0.65	-	1.29
Glycosphingolipid biosynthesis - ganglio series	6.67	-	13.33	-	-	-
TGF-beta signaling pathway	3.37	3.37	2.25	-	1.12	-
Taurine and hypotaurine metabolism	-	1-	-	-	-	-
Glycosaminoglycan degradation	4.76	-	4.76	-	-	-
Cysteine and methionine metabolism	2.94	2.94	-	-	-	-
Phenylalanine metabolism	5.88	-	5.88	-	-	-
Selenoamino acid metabolism	-	3.85	-	-	-	-
Glycosaminoglycan biosynthesis - keratan sulfate	6.67	-	6.67	-	-	-
Folate biosynthesis	-	9.09	-	-	-	-
D-Glutamine and D-glutamate metabolism	-	-	25.00	-	-	-
Glycosaminoglycan biosynthesis - heparan sulfate	-	-	3.85	-	-	-
Vitamin B6 metabolism	-	-	16.67	-	-	-

Table 12: Full *GO_HyperG* results.

GOBPID	Pvalue	OddsRatio	ExpCount	Count	Size	Term
GO:0006695	0.000564325	4.039671683	3.652889457	11	33	cholesterol biosynthetic process
GO:0019637	0.000581941	1.928817584	21.36386864	37	193	organophosphate metabolic process
GO:0044238	0.000617484	1.199125313	794.3374164	853	7176	primary metabolic process
GO:0044282	0.000669365	1.843316923	24.57398362	41	222	small molecule catabolic process
GO:0061035	0.000931469	8.06271302	1.328323439	6	12	regulation of cartilage development
GO:0045007	0.001353882	Inf	0.33208086	3	3	depurination
GO:0006006	0.00175041	1.931821932	17.2682047	30	156	glucose metabolic process
GO:0006694	0.00176201	2.199981282	11.40144285	22	103	steroid biosynthetic process
GO:0051346	0.001879224	3.127243823	4.759825656	12	43	negative regulation of hydrolase activity
GO:0046395	0.002250175	2.068631338	13.06184715	24	118	carboxylic acid catabolic process
GO:0048519	0.002282815	1.250419397	196.9239498	233	1779	negative regulation of biological process
GO:0055114	0.002517021	1.419127258	66.63755918	89	602	oxidation reduction
GO:0006986	0.002591974	2.580115577	6.863004434	15	62	response to unfolded protein
GO:0040007	0.002662201	1.478176221	51.25114601	71	463	growth
GO:0005996	0.00296543	1.703153822	24.90606448	39	225	monosaccharide metabolic process
GO:0051270	0.00344003	1.757315546	21.1424814	34	191	regulation of cellular component movement
GO:0006066	0.003585162	1.706666774	23.56749732	37	215	alcohol metabolic process
GO:0007050	0.003702983	2.047212782	12.06560457	22	109	cell cycle arrest
GO:0032332	0.003971791	10.73837077	0.774855339	4	7	positive regulation of chondrocyte differentiation
GO:0042127	0.004440273	1.347353547	83.02021492	106	750	regulation of cell proliferation
GO:0008361	0.004582991	1.609780542	28.78034117	43	260	regulation of cell size
GO:0000122	0.004709697	1.616759566	28.00548583	42	253	negative regulation of transcription from RNA polymerase II promoter
GO:0031324	0.005284238	1.349457597	78.14969565	100	706	negative regulation of cellular metabolic process
GO:0007017	0.005814032	1.593683262	28.33756669	42	256	microtubule-based process
GO:0006633	0.005848727	2.076657125	10.29450665	19	93	fatty acid biosynthetic process
GO:0043603	0.006022723	2.499934768	6.088149094	13	55	cellular amide metabolic process
GO:0008654	0.006261784	3.170616966	3.528700906	9	32	phospholipid biosynthetic process
GO:0006613	0.006700866	5.754671078	1.328323439	5	12	cotranslational protein targeting to membrane
GO:0050650	0.006700866	5.754671078	1.328323439	5	12	chondroitin sulfate proteoglycan biosynthetic process
GO:0048514	0.006872052	1.58480979	27.78409859	41	251	blood vessel morphogenesis
GO:0016481	0.006893935	1.424735097	49.81212895	67	450	negative regulation of transcription
GO:0045944	0.007028283	1.491020518	37.85721801	53	342	positive regulation of transcription from RNA polymerase II promoter
GO:0001709	0.007131938	3.395724807	2.988727737	8	27	cell fate determination

GO:0043436	0.00742927	1.370596693	63.09536334	82	570	oxoacid metabolic process
GO:0016043	0.007495528	1.182832729	284.3719095	320	2569	cellular component organization
GO:0001892	0.007639967	3.762301046	2.435259638	7	22	embryonic placenta development
GO:0048255	0.007665514	4.395984384	1.881791538	6	17	mRNA stabilization
GO:0048522	0.007902306	1.288618485	101.8767255	125	932	positive regulation of cellular process
GO:0045839	0.009908588	5.034911444	1.439017059	5	13	negative regulation of mitosis
GO:0042364	0.009952832	3.526858799	2.545953258	7	23	water-soluble vitamin biosynthetic process
GO:0002062	0.01033003	4.039302802	1.988120301	6	18	chondrocyte differentiation
GO:0009108	0.010501713	2.216963002	7.195085293	14	65	coenzyme biosynthetic process
GO:0051254	0.010623261	1.394943649	49.92282257	66	451	positive regulation of RNA metabolic process
GO:0010985	0.011394012	12.07346939	0.553468099	3	5	negative regulation of lipoprotein particle clearance
GO:0034382	0.011394012	12.07346939	0.553468099	3	5	chylomicron remnant clearance
GO:0042026	0.011394012	12.07346939	0.553468099	3	5	protein refolding
GO:0070071	0.011394012	12.07346939	0.553468099	3	5	proton-transporting two-sector ATPase complex assembly
GO:0009890	0.011411925	1.3467574	62.43120162	80	564	negative regulation of biosynthetic process
GO:0009083	0.011869143	6.450170416	0.995114251	4	9	branched chain family amino acid catabolic process
GO:0051172	0.011897624	1.363129096	56.34305253	73	509	negative regulation of nitrogen compound metabolic process
GO:0010941	0.011974357	1.284571514	90.43668746	111	817	regulation of cell death
GO:0000320	0.012245679	Inf	0.22138724	2	2	re-entry into mitotic cell cycle
GO:0001808	0.012245679	Inf	0.22138724	2	2	negative regulation of type IV hypersensitivity
GO:0002086	0.012245679	Inf	0.22138724	2	2	diaphragm contraction
GO:0002865	0.012245679	Inf	0.22138724	2	2	negative regulation of acute inflammatory response to antigenic stimulus
GO:0006436	0.012245679	Inf	0.22138724	2	2	tryptophanyl-tRNA aminoacylation
GO:0009162	0.012245679	Inf	0.22138724	2	2	deoxyribonucleoside monophosphate metabolic process
GO:0009211	0.012245679	Inf	0.22138724	2	2	pyrimidine deoxyribonucleoside triphosphate metabolic process
GO:0032753	0.012245679	Inf	0.22138724	2	2	positive regulation of interleukin-4 production
GO:0033615	0.012245679	Inf	0.22138724	2	2	mitochondrial proton-transporting ATP synthase complex assembly
GO:0034638	0.012245679	Inf	0.22138724	2	2	phosphatidylcholine catabolic process
GO:0043117	0.012245679	Inf	0.22138724	2	2	positive regulation of vascular permeability
GO:0045743	0.012245679	Inf	0.22138724	2	2	positive regulation of fibroblast growth factor receptor signaling pathway
GO:0050882	0.012245679	Inf	0.22138724	2	2	voluntary musculoskeletal movement
GO:0051365	0.012245679	Inf	0.22138724	2	2	cellular response to potassium ion starvation

GO:0060696	0.012245679	Inf	0.22138724	2	2	regulation of phospholipid catabolic process
GO:0090042	0.012245679	Inf	0.22138724	2	2	tubulin deacetylation
GO:0007040	0.012736673	3.319115641	2.656646878	7	24	lysosome organization
GO:0010551	0.012792657	1.896063647	11.06936199	19	100	regulation of gene-specific transcription from RNA polymerase II promoter
GO:0009966	0.012986139	1.270704802	97.07830465	118	877	regulation of signal transduction
GO:0045017	0.01381122	4.494463536	1.543806647	5	14	glycerolipid biosynthetic process
GO:0001569	0.013855587	3.719049866	2.103178778	6	19	patterning of blood vessels
GO:0046460	0.01403032	4.475098395	1.549710679	5	14	neutral lipid biosynthetic process
GO:0006638	0.014189586	2.695453507	3.975626269	9	36	neutral lipid metabolic process
GO:0006662	0.014196153	2.69522504	3.975925369	9	36	glycerol ether metabolic process
GO:0019362	0.014355365	2.520933014	4.649132036	10	42	pyridine nucleotide metabolic process
GO:0031667	0.014411399	1.619404019	19.92485158	30	180	response to nutrient levels
GO:0009101	0.0151723	1.64228474	18.3751409	28	166	glycoprotein biosynthetic process
GO:0019320	0.015523882	2.093265301	7.527166153	14	68	hexose catabolic process
GO:0006284	0.016038804	3.134455055	2.767340497	7	25	base-excision repair
GO:0044264	0.016671005	2.141529214	6.863004434	13	62	cellular polysaccharide metabolic process
GO:0009394	0.017980316	3.453111306	2.213872398	6	20	2'-deoxyribonucleotide metabolic process
GO:0030968	0.017980316	3.453111306	2.213872398	6	20	endoplasmic reticulum unfolded protein response
GO:0006098	0.018142932	5.367823916	1.106936199	4	10	pentose-phosphate shunt
GO:0007094	0.018142932	5.367823916	1.106936199	4	10	mitotic cell cycle spindle assembly checkpoint
GO:0001944	0.018926412	1.45505431	31.28129935	43	283	vasculature development
GO:0042981	0.018996923	1.263540134	88.31052632	107	799	regulation of apoptosis
GO:0007052	0.019163493	4.027247956	1.660404298	5	15	mitotic spindle organization
GO:0046504	0.019163493	4.027247956	1.660404298	5	15	glycerol ether biosynthetic process
GO:0001558	0.019891772	1.558665932	21.25317502	31	192	regulation of cell growth
GO:0006493	0.019904143	2.969232426	2.878034117	7	26	protein amino acid O-linked glycosylation
GO:0043487	0.019904143	2.969232426	2.878034117	7	26	regulation of RNA stability
GO:0032787	0.020523893	1.415839234	35.75403923	48	323	monocarboxylic acid metabolic process
GO:0016265	0.020527368	1.218333815	125.526565	147	1134	death
GO:0051235	0.020720544	1.900117082	9.298264072	16	84	maintenance of location
GO:0043068	0.020782179	1.357462455	47.15548208	61	426	positive regulation of programmed cell death
GO:0007089	0.020922336	8.04829932	0.664161719	3	6	traversing start control point of mitotic cell cycle
GO:0034379	0.020922336	8.04829932	0.664161719	3	6	very-low-density lipoprotein particle assembly
GO:0046325	0.020922336	8.04829932	0.664161719	3	6	negative regulation of glucose import
GO:0048193	0.021948448	1.693662827	14.05808973	22	127	Golgi vesicle transport

GO:0051186	0.022002926	1.556319633	20.5890133	30	186	cofactor metabolic process
GO:0001763	0.022708027	1.922235286	8.634102352	15	78	morphogenesis of a branching structure
GO:0009310	0.022708027	1.922235286	8.634102352	15	78	amine catabolic process
GO:0006099	0.022872745	3.22263122	2.324566018	6	21	tricarboxylic acid cycle
GO:0008633	0.022872745	3.22263122	2.324566018	6	21	activation of pro-apoptotic gene products
GO:0071445	0.022872745	3.22263122	2.324566018	6	21	cellular response to protein stimulus
GO:0009820	0.022972409	2.304267161	4.981212895	10	45	alkaloid metabolic process
GO:0071216	0.024374415	2.82053206	2.988727737	7	27	cellular response to biotic stimulus
GO:0032869	0.025260072	1.892039609	8.744795972	15	79	cellular response to insulin stimulus
GO:0046627	0.025390202	3.66082487	1.771097918	5	16	negative regulation of insulin receptor signaling pathway
GO:0030335	0.025449235	1.806372903	10.29450665	17	93	positive regulation of cell migration
GO:0016051	0.025831105	1.661722177	14.27947697	22	129	carbohydrate biosynthetic process
GO:0016053	0.026018048	1.571017045	18.3751409	27	166	organic acid biosynthetic process
GO:0006641	0.028205529	3.031335616	2.427883168	6	22	triglyceride metabolic process
GO:0044255	0.028339134	1.333356746	47.06397638	60	429	cellular lipid metabolic process
GO:0044262	0.02838336	1.586802395	16.85830207	25	154	cellular carbohydrate metabolic process
GO:0050879	0.028585037	3.020961145	2.435259638	6	22	multicellular organismal movement
GO:0060070	0.028585037	3.020961145	2.435259638	6	22	Wnt receptor signaling pathway through beta-catenin
GO:0008203	0.028607713	2.028275862	6.608407413	12	60	cholesterol metabolic process
GO:0010876	0.028815549	1.540550338	19.37138348	28	175	lipid localization
GO:0030178	0.029105987	2.480021003	3.763583077	8	34	negative regulation of Wnt receptor signaling pathway
GO:0048870	0.029378738	1.337046258	45.38438416	58	410	cell motility
GO:0046474	0.030403767	1.942415018	7.416472533	13	67	glycerophospholipid biosynthetic process
GO:0009119	0.030479016	2.179342706	5.202600135	10	47	ribonucleoside metabolic process
GO:0015980	0.030724272	1.590062112	16.16126851	24	146	energy derivation by oxidation of organic compounds
GO:0044275	0.030930064	1.79394494	9.741038551	16	88	cellular carbohydrate catabolic process
GO:0009891	0.031125489	1.260243047	72.61501465	88	656	positive regulation of biosynthetic process
GO:0031325	0.031128383	1.228889341	93.75749605	111	847	positive regulation of cellular metabolic process
GO:0042176	0.032178796	2.063134286	5.977455475	11	54	regulation of protein catabolic process
GO:0001578	0.032775489	3.355472298	1.881791538	5	17	microtubule bundle formation
GO:0006739	0.032775489	3.355472298	1.881791538	5	17	NADP metabolic process
GO:0046503	0.032775489	3.355472298	1.881791538	5	17	glycerolipid catabolic process
GO:0006882	0.033635244	6.035714286	0.774855339	3	7	cellular zinc ion homeostasis
GO:0006978	0.033635244	6.035714286	0.774855339	3	7	DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator
GO:0009103	0.033635244	6.035714286	0.774855339	3	7	lipopolysaccharide biosynthetic process

GO:0017085	0.033635244	6.035714286	0.774855339	3	7	response to insecticide
GO:0030219	0.033635244	6.035714286	0.774855339	3	7	megakaryocyte differentiation
GO:0031112	0.033635244	6.035714286	0.774855339	3	7	positive regulation of microtubule polymerization or depolymerization
GO:0033032	0.033635244	6.035714286	0.774855339	3	7	regulation of myeloid cell apoptosis
GO:0051006	0.033635244	6.035714286	0.774855339	3	7	positive regulation of lipoprotein lipase activity
GO:0035108	0.033978947	1.769219921	9.851732171	16	89	limb morphogenesis
GO:0006552	0.034029275	16.08837525	0.33208086	2	3	leucine catabolic process
GO:0007549	0.034029275	16.08837525	0.33208086	2	3	dosage compensation
GO:0008356	0.034029275	16.08837525	0.33208086	2	3	asymmetric cell division
GO:0009221	0.034029275	16.08837525	0.33208086	2	3	pyrimidine deoxyribonucleotide biosynthetic process
GO:0014012	0.034029275	16.08837525	0.33208086	2	3	axon regeneration in the peripheral nervous system
GO:0015808	0.034029275	16.08837525	0.33208086	2	3	L-alanine transport
GO:0015917	0.034029275	16.08837525	0.33208086	2	3	aminophospholipid transport
GO:0032727	0.034029275	16.08837525	0.33208086	2	3	positive regulation of interferon-alpha production
GO:0042149	0.034029275	16.08837525	0.33208086	2	3	cellular response to glucose starvation
GO:0043517	0.034029275	16.08837525	0.33208086	2	3	positive regulation of DNA damage response, signal transduction by p53 class mediator
GO:0051964	0.034029275	16.08837525	0.33208086	2	3	negative regulation of synaptogenesis
GO:0051988	0.034029275	16.08837525	0.33208086	2	3	regulation of attachment of spindle microtubules to kinetochore
GO:0070050	0.034029275	16.08837525	0.33208086	2	3	neuron homeostasis
GO:0070272	0.034029275	16.08837525	0.33208086	2	3	proton-transporting ATP synthase complex biogenesis
GO:0052547	0.034198223	1.806860758	9.076876832	15	82	regulation of peptidase activity
GO:0006213	0.035161109	2.843016961	2.545953258	6	23	pyrimidine nucleoside metabolic process
GO:0009954	0.035161109	2.843016961	2.545953258	6	23	proximal/distal pattern formation
GO:0050680	0.035161109	2.843016961	2.545953258	6	23	negative regulation of epithelial cell proliferation
GO:0006084	0.035276484	2.563685973	3.210114977	7	29	acetyl-CoA metabolic process
GO:0006984	0.035433483	4.036910458	1.324904042	4	12	ER-nuclear signaling pathway
GO:0019432	0.03573195	4.025187202	1.328323439	4	12	triglyceride biosynthetic process
GO:0008104	0.036311138	1.205928573	107.3728113	125	970	protein localization
GO:0007173	0.037429701	2.198397914	4.649132036	9	42	epidermal growth factor receptor signaling pathway
GO:0035239	0.039019155	1.60018712	14.05808973	21	127	tube morphogenesis
GO:0055086	0.039444074	1.316077036	44.38814158	56	401	nucleobase, nucleoside and nucleotide metabolic process
GO:0001764	0.039567165	2.067231015	5.423987375	10	49	neuron migration

GO:0006720	0.039567165	2.067231015	5.423987375	10	49	isoprenoid metabolic process
GO:0070887	0.039617233	1.370904686	32.87600511	43	297	cellular response to chemical stimulus
GO:0044237	0.04033255	1.117182034	502.5485581	531	4814	cellular metabolic process
GO:0045941	0.040709611	1.272308874	58.00345683	71	524	positive regulation of transcription
GO:0000079	0.040882124	1.971105031	6.198842714	11	56	regulation of cyclin-dependent protein kinase activity
GO:0051225	0.041366394	3.097097045	1.992485158	5	18	spindle assembly
GO:0007224	0.041769828	2.452013761	3.320808597	7	30	smoothened signaling pathway
GO:0042246	0.041769828	2.452013761	3.320808597	7	30	tissue regeneration
GO:0006829	0.042635921	2.684844354	2.656646878	6	24	zinc ion transport
GO:0009109	0.042635921	2.684844354	2.656646878	6	24	coenzyme catabolic process
GO:0045669	0.042635921	2.684844354	2.656646878	6	24	positive regulation of osteoblast differentiation
GO:0007049	0.044211846	1.201327434	100.7311941	117	910	cell cycle
GO:0051253	0.045363134	1.323724618	39.40692868	50	356	negative regulation of RNA metabolic process
GO:0042476	0.045792608	1.928091358	6.309536334	11	57	odontogenesis
GO:0006766	0.045847849	1.764693283	8.634102352	14	78	vitamin metabolic process
GO:0008624	0.04598303	1.632572686	11.84421733	18	107	induction of apoptosis by extracellular signals
GO:0016578	0.047208999	3.577641631	1.439017059	4	13	histone deubiquitination
GO:0019433	0.047208999	3.577641631	1.439017059	4	13	triglyceride catabolic process
GO:0031577	0.047208999	3.577641631	1.439017059	4	13	spindle checkpoint
GO:0043154	0.047208999	3.577641631	1.439017059	4	13	negative regulation of caspase activity
GO:0043193	0.04838207	1.676741927	10.29450665	16	93	positive regulation of gene-specific transcription
GO:0048736	0.04838207	1.676741927	10.29450665	16	93	appendage development
GO:0010827	0.048990293	2.349647567	3.431502217	7	31	regulation of glucose transport
GO:0051650	0.048990293	2.349647567	3.431502217	7	31	establishment of vesicle localization
GO:0007010	0.049314479	1.272046647	52.24738859	64	472	cytoskeleton organization
GO:0001945	0.049466161	4.828163265	0.885548959	3	8	lymph vessel development
GO:0002262	0.049466161	4.828163265	0.885548959	3	8	myeloid cell homeostasis
GO:0008354	0.049466161	4.828163265	0.885548959	3	8	germ cell migration
GO:0009113	0.049466161	4.828163265	0.885548959	3	8	purine base biosynthetic process
GO:0009263	0.049466161	4.828163265	0.885548959	3	8	deoxyribonucleotide biosynthetic process
GO:0018065	0.049466161	4.828163265	0.885548959	3	8	protein-cofactor linkage
GO:0030858	0.049466161	4.828163265	0.885548959	3	8	positive regulation of epithelial cell differentiation
GO:0035313	0.049466161	4.828163265	0.885548959	3	8	wound healing, spreading of epidermal cells
GO:0043094	0.049466161	4.828163265	0.885548959	3	8	cellular metabolic compound salvage

QualCont FUNCTION

```
QualCont<-function (FilesPath,NormType,pdfile)
{
#FilesPath: location of the .CEL files
#NormType: string defining the type of normalization method; "rma", "gcrma" or
#            "mas5"
#pdfile: phenodata file created by user; tab delimited .txt file

#rma and gcrma generates logged values of signal intensity while mas5 not..
#therefore log status of the signal intensity is set to 1 or 0.
#This information is needed for log fold-change calculation in DEGidentifier
#function...
if(NormType=="mas5") LOG.STAT<-0 else LOG.STAT<-1
assign("LOG.STAT", LOG.STAT, envir=.GlobalEnv)

library(affy)
library(affyPLM)

#FilesPath given by the user is set as the
#working directory.

setwd(FilesPath)


#phenodata file is read;
#.CEL files are read with phenodata file...

read.AnnotatedDataFrame(pdfile,header=TRUE,row.names=1,as.is=TRUE)->pd
assign("pd",pd, envir=.GlobalEnv)
rawAffyData <- ReadAffy(filenames=pData(pd)$FileName,phenoData=pd,verbose=TRUE)
assign("rawAffyData",rawAffyData,envir=.GlobalEnv)
groups<-levels(as.factor(pData(pd)[,2]))
assign("groups", groups, envir=.GlobalEnv)
ind=vector(mode="integer",length=length(pData(pd)[,2]))

if (length(groups)<2) {
cat ("There must be at least 2 groups defined")
break
}

for (i in 1:length(groups)) {
tempind=which(pData(pd)[,2]==groups[i])
ind[tempind]<-rep(i,length(tempind))
}

assign("ind",ind, envir=.GlobalEnv)


#normalization is performed...

switch (NormType,
rma=assign("normdata", rma(rawAffyData)),
mas5=assign("normdata", mas5(rawAffyData)),
gcrma=assign("normdata", gcrma(rawAffyData))
)

if(exists("normdata")) {
ndata<-exprs(normdata)

#normalized data matrix and groups are saved into
#Global environment for further and easier usage in
#other functions... see DEGidentifier and Grapher...
```

```
assign("ndata",ndata, envir=.GlobalEnv)
assign("groups",groups, envir=.GlobalEnv)
}
else
{
paste("This Function does not know a normalization method
called",NormType,"\n")->err1
cat(err1)
break
}

#normalized data is saved into working directory...
write(ndata, file="ndata.xls", sep="\t")

#.CEL file names are arranged for output file denomination...
#last 4 characters are deleted (.CEL)
colnames(ndata)->cfn
cfn<-as.vector(sapply(cfn, function(x) substr(x,1,(nchar(x)-4))))

#chip images are generated...
for (i in 1:ncol(ndata)){
jpeg(paste(cfn[i],".jpeg",sep=""),width=1080,height=840)
image(rawAffyData[,i])
dev.off()
}

#Histograms of in-array signal intensity distribution
#of raw data for each array are generated...
for(i in 1:ncol(ndata)){
jpeg(file=paste(cfn[i],"HIST_raw.jpeg",sep=""),width=1080,height=840)
hist(log2(intensity(rawAffyData[,i])),breaks = 100, col =
"blue",main=paste("Histogram of raw", cfn[i]),xlab=cfn[i])
dev.off()
}


#Histograms of in-array signal intensity distribution
#of normalized data for each array are generated...
for(i in 1:ncol(ndata)){
jpeg(file=paste(cfn[i],"HIST_normalized.pjpeg",sep=""),width=1080,height=840)
hist(ndata[,i],breaks = 100, col = "blue",main=paste("Histogram of normalized",
cfn[i]),xlab=cfn[i])
dev.off()
}

#Boxplots of in-array signal intensity distribution
#of raw data for each array are generated...
jpeg(file="raw_BOXPLOT.jpeg",width=1080,height=840)
boxplot(rawAffyData, col="red", main="Boxplot of Raw Data", xlab="Chip Names",
ylab=
"Raw Intensity Values")
dev.off()


#Boxplots of in-array signal intensity distribution
#of normalized data for each array are generated...
jpeg(file="normalized_BOXPLOT.jpeg",width=1080,height=840)
boxplot(ndata, col="red", main="Boxplot of Normalized Data", xlab="Chip Names",
ylab= "Normalized Intensity Values")
dev.off()
```

```
#paired MA PLOTs of the raw data are generated...
jpeg(file="MAPlots_raw.jpeg",width=1080,height=840)
mva.pairs(pm(rawAffyData))
dev.off()

#paired MA PLOTs of the normalized data are generated...
pdf(file="MAPlots_norm.pdf",width=1080,height=840)
mva.pairs(ndata)
dev.off()

#RNA quality control assessments....
jpeg(file="RNA_Degradation_Plot.jpeg",width=1080,height=840)
RNA_degradation<-AffyRNAdeg(rawAffyData,log.it=TRUE)
summary_RNADEG<-summaryAffyRNAdeg(RNA_degradation,signif.digits=3)
write.table(summary_RNADEG, file="summaryRNADEG.xls", sep="\t")
plotAffyRNAdeg(RNA_degradation, transform = "shift.scale", cols =
rainbow(ncol(ndata)))
dev.off()

#RLE (Relative Log Expression) plot is generated...
Pset<-fitPLM(rawAffyData)
jpeg(file="RLE_BoxPlot.jpeg",width=1080,height=840)
RLE(Pset,main="RLE for Raw Data")
dev.off()

}
#End of the function....
```

DEGidentifier FUNCTION

```
DEGidentifier<-function(ndata,pdfile,lfc,method,adjust.method,p.value){

#ndata:       normalized data matrix; exprs() funtion output...
#pdfile:      experiment descriptor (phenodata) file...
#lfc:         log fold change treshold to be applied for filtering...
#method:      character string specify how probes and contrasts are to be
#             combined in the multiple testing strategy.  Choices are
#             '"separate"', '"global"', '"hierarchical"',
#             '"nestedF"' or any partial string.
#adjust.method:     character string specifying p-value adjustment method.
#                   Possible values are '"none"', '"BH"', '"fdr"'
#                   (equivalent to '"BH"'), '"BY"' and '"holm"'.
#p.value:     numeric value between 0 and 1 giving the desired pvalue
#             treshold to be applied for the test...


method <- match.arg(method, c("separate", "global", "hierarchical","nestedF"))
adjust.method <- match.arg(adjust.method, c("none", "bonferroni","holm", "BH",
"fdr", "BY"))


library(limma)
library(hgu133plus2.db)
library(made4)
#made4 package is dependent on the following packages:
#ade4,gplots,gtools,gdata,caTools,bitops and scatterplot...
#ensure all these packages are already installed before calling made4...


#Generation of contrast matrix from phenodata file automatically...
pd<-readTargets(pdfile,sep="")
f<-paste(pd$Target,sep="")
f<-factor(f)

design<-model.matrix(~0+f)
colnames(design)<-levels(f)

ContrastTemplet<-vector()

TempComb<-combn(colnames(design),2)
for (i in 1:ncol(TempComb)){
ContrastTemplet[i]<-paste(TempComb[2,i],"-",TempComb[1,i],sep="")
}
ContrastTemplet[1]->cd
assign("cd",cd,envir=.GlobalEnv)
contrast.matrix<-makeContrasts(cd,levels=design)

if (length(levels(f)>=3){
for (i in 2:length(ContrastTemplet)){
ContrastTemplet[i]->cd
assign("cd",cd,envir=.GlobalEnv)
Temp.contrast.matrix<-makeContrasts(cd,levels=design)
contrast.matrix<-cbind(contrast.matrix,Temp.contrast.matrix)
}
}

#contrast matrix is saved to global environment to enable user
#to check whether it is correct...
assign("contrast.matrix", contrast.matrix,envir=.GlobalEnv)
```

```
#fitting the data to a linear model and
#emprical Bayes moderation....
fit<-lmFit(ndata,design)
fit2<-contrasts.fit(fit,contrast.matrix)
fit2<-eBayes(fit2)
assign("fit",fit,envir=.GlobalEnv)
assign("fit2", fit2, envir=.GlobalEnv)

#Top 10 testing results for each comparison to give an idea...
print("These are some of your top results...")
for(i in 1:length(ContrastTemplet)){
print(ContrastTemplet[i])
print(topTable(fit2,coef=i,adjust="BH"),zero.print=".")
}


#Actual testing and determination of DEGs with parameters given by the user...
results<-
decideTests(fit2,method=method,adjust.method=adjust.method,p.value=p.value,lfc=
lfc)

#testing results are saved to Global environment to enable
#user to check and for further and easier usage in other functions,
#see KEGGscript and Grapher...
assign("results",results,envir=.GlobalEnv)

#Generation of Venn Diagram for the DEGs....
jpeg(file="VennDiagram.pdf",width=1080,height=840)
vennDiagram(results)
dev.off()

#List of all significant probes is generated and saved to working directory...
TempSignif<-vector()
TempSignif<-which(results[,1]!=0)
TempSignifProbes<-names(TempSignif)
write.table(TempSignifProbes,file=paste(colnames(results)[1],"SignifDEGProbes.x
ls",sep=""),sep="\t")
AllSignifProbes<-TempSignifProbes

if (length(levels(f)>=3){
for(i in 2:length(colnames(results))){
TempSignif<-which(results[,i]!=0)
TempSignifProbes<-names(TempSignif)
write.table(TempSignifProbes,file=paste(colnames(results)[i],"SignifDEGProbes.x
ls",sep=""),sep="\t")
AllSignifProbes<-union(AllSignifProbes,names(TempSignif))
}
}

write.table(AllSignifProbes,file="AllSignifDEGProbes.xls",sep="\t")
assign("AllSignifProbes",AllSignifProbes,envir=.GlobalEnv)


#Lists of significant up- and downregulated probes and gene symbols
#for each comparison are generated and saved to working directory...
UPs<-vector()
for(i in 1:ncol(results)){
UPs<-which(results[,i]==1)
UPsProbes<-names(UPs)
UPsGenes<-unlist(mget(UPsProbes,hgu133plus2SYMBOL,ifnotfound=NA))
UPsGenes<-as.vector(UPsGenes)
UPsGenes<-UPsGenes[!is.na(UPsGenes)]
UPsGenes<-UPsGenes[!duplicated(UPsGenes)]
```

```
write.table(UPsProbes,file=paste(colnames(results)[i],"UPsProbes.xls",sep=""),s
ep="\t")
write.table(UPsGenes,file=paste(colnames(results)[i],"UPsGenes.xls",sep=""),sep
="\t")
}



DOWNs<-vector()
for (i in 1:ncol(results)){
DOWNs<-which(results[,i]==-1)
DOWNsProbes<-names(DOWNs)
DOWNsGenes<-unlist(mget(DOWNsProbes,hgu133plus2SYMBOL,ifnotfound=NA))
DOWNsGenes<-as.vector(DOWNsGenes)
DOWNsGenes<-DOWNsGenes[!is.na(DOWNsGenes)]
DOWNsGenes<-DOWNsGenes[!duplicated(DOWNsGenes)]
write.table(DOWNsProbes,file=paste(colnames(results)[i],"DOWNsProbes.xls",sep="
"),sep="\t")
write.table(DOWNsGenes,file=paste(colnames(results)[i],"DOWNsGenes.xls",sep="")
,sep="\t")
}


#normalized data matrix of significant probes,"ndata2",
#is extracted from whole normalized data matrix, "ndata", and
#saved into global environment for further usage (see heatmap generation below)
#and saved into working directory...
Selected<-charmatch(rownames(ndata),AllSignifProbes)
Indxs<-which(Selected!="NA")

ndata2<-ndata[Indxs,]

assign("ndata2",ndata2,envir=.GlobalEnv)
write.table(ndata2,file="ndata.Of.Signif.Probes.txt",sep="\t")


#HEATMAP GENERATION
jpeg(file="HeatMap.jpeg",width=1080,height=840)
heatplot(ndata2)
dev.off()



#averages of multiple replicates are calculated for log fold-change
calculation...
tempData<-ndata2
colnames(tempData)<-as.vector(f)
index<-which(colnames(tempData)==levels(f)[1])
meanGroups<-rowMeans(tempData[,index[1]:index[length(index)]])
meanGroups<-as.data.frame(meanGroups)


for (i in 2:length(levels(f))){
index<-which(colnames(tempData)==levels(f)[i])
TempMeanGroups<-rowMeans(tempData[,index[1]:index[length(index)]])
TempmeanGroups<-as.data.frame(TempMeanGroups)
meanGroups<-cbind(meanGroups,TempMeanGroups)
}
colnames(meanGroups)<-levels(f)


#Log fold-change values are generated and saved into global environment
#for further and easier usage by other functions (see Grapher),
```

```
#and saved into working directory for possible further usage by user...
fc<-matrix(0,nrow=nrow(meanGroups),ncol=ncol(combn(ncol(meanGroups),2)))
rownames(fc)<-rownames(meanGroups)
colnames(fc)<-vector(mode="list",length=ncol(meanGroups))
meanGroups<-as.matrix(meanGroups)

TempNames<-vector()
Tempfc<-matrix()
p=1

if(LOG.STAT==1){
for(i in 1:(ncol(meanGroups)-1)){
for(j in (i+1):ncol(meanGroups)){
Tempfc<-meanGroups[,j]-meanGroups[,i]
TempNames<-paste(colnames(meanGroups)[j],"-",colnames(meanGroups)[i],sep="")
fc[,p]<-Tempfc
colnames(fc)[p]<-TempNames
p=p+1
}
}
}

else{
for(i in 1:(ncol(meanGroups)-1)){
for(j in (i+1):ncol(meanGroups)){
Tempfc<-meanGroups[,j]/meanGroups[,i]
TempNames<-paste(colnames(meanGroups)[j],"-",colnames(meanGroups)[i],sep="")
fc[,p]<-Tempfc
colnames(fc)[p]<-TempNames
p=p+1
}
}
}

assign("fc",fc,envir=.GlobalEnv)
write.table(fc,file="fc_all.xls", sep="\t")


#Plots for fold-change correspondences between comparisons are generated
#and saved into working directory...

Pairs<-combn(colnames(fc),2)

for(i in 1:ncol(Pairs)){
Corresp.Title=paste("FC Corresp.", Pairs[1,i], "vs", Pairs[2,i], sep=" ")
jpeg(file=paste(Corresp.Title, ".jpeg", sep=""))
plot(fc[,Pairs[1,i]],fc[,Pairs[2,i]], main=Corresp.Title, xlab=Pairs[1,i],
ylab=Pairs[2,i])
dev.off()
}

}
#End of the function...
```

KEGGscript FUNCTION

```
KEGGscript<-function(results){

require(hgu133plus2.db)
require(KEGG.db)
require(KEGGgraph)
require(Rgraphviz)
require(GeneAnswers)

####KEGG Pathway Annotation######

kegg_Paths<-mget(AllSignifProbes,hgu133plus2PATH, ifnotfound=NA)
listem<-t(as.data.frame(kegg_Paths[1]))[1,]
listem<-as.vector(listem)
for(i in 2:length(kegg_Paths)){
listem<-union(listem,t(as.data.frame(kegg_Paths[i]))[1,])
}

listem<-listem[which(listem!="NA")]
assign("KEGGids",listem,envir=.GlobalEnv)
write.table(listem, file="KEGGids.txt",sep="\t")
##########################################################
##########################################################

#download the KGML files...
destdir<-system.file("extdata",package="KEGGgraph",lib.loc=NULL)

#for(i in 1:length(KEGGids)){
#retrieveKGML(KEGGids[i],organism="hsa",destfile=paste(destdir,"\\hsa",KEGGids[
i],".xml",sep=""),method="internal",quiet=TRUE)
#}

#read KGML files in and parse them as graph objects.
FileList<-list()
       for (i in 1:length(KEGGids)){
              FileList[i]<-
system.file(paste("extdata/hsa",KEGGids[i],".xml",sep=""),package="KEGGgraph")
       }
FileList<-FileList[is.na(charmatch(FileList,""))]
FileList<-unlist(FileList)

###KEGG ANNOTATION OF UP- AND DOWNREGULATED GENES IN EACH GROUP
EntrezIDMaP <- hgu133plus2ENTREZID
       for(i in 1:ncol(results)){
              write(paste(colnames(results)[i],"KEGG PATHWAY
RESULTS...",sep="\n"),file=paste(colnames(results)[i],"KEGG_RESULTS.txt",sep=""
),                  ncolumns=if(is.character(paste(colnames(results)[i],"KEGG
PATHWAY RESULTS...",sep="\n"))) 1 else 5,append=FALSE)

              RESTAB<-matrix(0,ncol=2,nrow=length(FileList))
              colnames(RESTAB)<-c("UP Percentage","DOWN Percentage")
              rownames(RESTAB)<-vector(mode="list",length=length(FileList))


              for(j in 1:length(FileList)){
                     KEGG_PATH<-parseKGML(FileList[j])
                     KEGG_GRAPH<-KEGGpathway2Graph(KEGG_PATH,expandGenes=TRUE)


                     UPs<-vector()
```

```
                    UPs<-which(results[,i]==1)
                    UPsENTREZ<-unlist(as.list(EntrezIDMaP[names(UPs)]))
                    UPsENTREZ<-as.vector(UPsENTREZ)
                    UPsENTREZ<-UPsENTREZ[!is.na(UPsENTREZ)]
                    UPsENTREZ<-UPsENTREZ[!duplicated(UPsENTREZ)]
                    UPsKEGGID<-translateGeneID2KEGGID(UPsENTREZ)

                    DOWNs<-vector()

                    DOWNs<-which(results[,i]==-1)
                    DOWNsENTREZ<-unlist(as.list(EntrezIDMaP[names(DOWNs)]))
                    DOWNsENTREZ<-as.vector(DOWNsENTREZ)
                    DOWNsENTREZ<-DOWNsENTREZ[!is.na(DOWNsENTREZ)]
                    DOWNsENTREZ<-DOWNsENTREZ[!duplicated(DOWNsENTREZ)]
                    DOWNsKEGGID<-translateGeneID2KEGGID(DOWNsENTREZ)


                    isMappedUP<-nodes(KEGG_GRAPH)%in% UPsKEGGID
                    if(mean(isMappedUP)==0){
                    declarUP<-sprintf("No genes are upregulated...")
                    }
                    else{
                    declarUP<-sprintf("%2.2f%% of genes are
upregulated...",mean(isMappedUP)*100)
                    declarUP<-
paste("************************","************************",KEGG_PATH@pathway
Info@title, declarUP, sep="\n")
                    declarUP<-paste(declarUP,"GENES ARE:",sep="\n")

                    GENSYMBOLUP<-
unlist(mget(translateKEGGID2GeneID(nodes(KEGG_GRAPH)[isMappedUP==TRUE]),
org.Hs.egSYMBOL, ifnotfound=NA))


      write(declarUP,file=paste(colnames(results)[i],"KEGG_RESULTS.txt",sep=""
), ncolumns=if(is.character(declarUP)) 1 else 5, append=TRUE)

      write(GENSYMBOLUP,file=paste(colnames(results)[i],"KEGG_RESULTS.txt",sep
=""), ncolumns=if(is.character(GENSYMBOLUP)) 1 else 5, append=TRUE)
                    }

                    isMappedDOWN<-nodes(KEGG_GRAPH)%in% DOWNsKEGGID
                    if(mean(isMappedDOWN)==0){
                    declarDOWN<-sprintf("No genes are downregulated...")
                    }
                    else{
                    declarDOWN<-sprintf("%2.2f%% of genes are
downregulated...",mean(isMappedDOWN)*100)
                    declarDOWN<-
paste("************************","************************",
KEGG_PATH@pathwayInfo@title, declarDOWN, sep="\n")
                    declarDOWN<-paste(declarDOWN,"GENES ARE:",sep="\n")

                    GENSYMBOLDOWN<-
unlist(mget(translateKEGGID2GeneID(nodes(KEGG_GRAPH)[isMappedDOWN==TRUE]),
org.Hs.egSYMBOL, ifnotfound=NA))
                    }

      write(declarDOWN,file=paste(colnames(results)[i],"KEGG_RESULTS.txt",sep=
""), ncolumns=if(is.character(declarDOWN)) 1 else 5, append=TRUE)

      write(GENSYMBOLDOWN,file=paste(colnames(results)[i],"KEGG_RESULTS.txt",s
ep=""), ncolumns=if(is.character(GENSYMBOLDOWN)) 1 else 5, append=TRUE)
```

```
                    rownames(RESTAB)[j]<-KEGG_PATH@pathwayInfo@title
                    RESTAB[j,1]<-mean(isMappedUP)*100
                    RESTAB[j,2]<-mean(isMappedDOWN)*100
                    }


      write.table(RESTAB,file=paste("KEGG_TABLE",colnames(results)[i],".xls",s
ep=""),sep="\t")


x<-
geneAnswersBuilder(UPsENTREZ,"org.Hs.eg.db",categoryType="GO.BP",testType="hype
rG",pvalueT=0.1,FDR.correction=TRUE,geneExpressionProfile=NULL)
y<-
geneAnswersBuilder(UPsENTREZ,"org.Hs.eg.db",categoryType="KEGG",testType="hyper
G",pvalueT=0.1,FDR.correction=TRUE,geneExpressionProfile=NULL)
z<-
geneAnswersBuilder(DOWNsENTREZ,"org.Hs.eg.db",categoryType="GO.BP",testType="hy
perG",pvalueT=0.1,FDR.correction=TRUE,geneExpressionProfile=NULL)
w<-
geneAnswersBuilder(DOWNsENTREZ,"org.Hs.eg.db",categoryType="KEGG",testType="hyp
erG",pvalueT=0.1,FDR.correction=TRUE,geneExpressionProfile=NULL)


            xx<-geneAnswersReadable(x)
            yy<-geneAnswersReadable(y)
            zz<-geneAnswersReadable(z)
            ww<-geneAnswersReadable(w)



            #if chartType in "chartPlotter" is given as "all", it opens
            #two devices and this raises problems in drawing. this loop
            #is written for this problem...

            TYPE_CHART<-c("pieChart","barPlot")

            for (m in 1:length(TYPE_CHART)){

                  if(length(rownames(xx@enrichmentInfo))>1){

                  pdf(file=paste(colnames(results)[i],
"_UPREG_GO.BP_TOP5_",TYPE_CHART[m],".pdf",sep=""),width=14,height=9)
                  ChartPlotter(xx, chartType=TYPE_CHART[m])
                  dev.off()
                  }
                  else{print(paste("The only term: ",xx@enrichmentInfo,
sep="\n"))}

                  if(length(rownames(yy@enrichmentInfo))>1){
                  pdf(file=paste(colnames(results)[i],
"_UPREG_KEGG_TOP5_",TYPE_CHART[m],".pdf",sep=""),width=14,height=9)
                  ChartPlotter(yy, chartType=TYPE_CHART[m])
                  dev.off()
                  }

                  else{print(paste("The only term: ",yy@enrichmentInfo,
sep="\n"))}


                  if(length(rownames(zz@enrichmentInfo))>1){
```

```
                    pdf(file=paste(colnames(results)[i],
"_DOWNREG_GO.BP_TOP5_",TYPE_CHART[m],".pdf",sep=""),width=14,height=9)
                    ChartPlotter(zz, chartType=TYPE_CHART[m])
                    dev.off()
                    }
                    else{print(paste("The only term: ",zz@enrichmentInfo,
sep="\n"))}

                    if(length(rownames(ww@enrichmentInfo))>1){
                    pdf(file=paste(colnames(results)[i],
"_DOWNREG_KEGG_TOP5_",TYPE_CHART[m],".pdf",sep=""),width=14,height=9)
                    ChartPlotter(ww, chartType=TYPE_CHART[m])
                    dev.off()
                    }
                    else{print(paste("The only term: ",ww@enrichmentInfo,
sep="\n"))}
             }

       }


AllEntrez<-unlist(as.list(EntrezIDMaP[AllSignifProbes]))
AllEntrez<-AllEntrez[!is.na(AllEntrez)]

xAll<-
geneAnswersBuilder(AllEntrez,"org.Hs.eg.db",categoryType="GO.BP",testType="hype
rG", pvalueT=0.1, FDR.correction=TRUE,geneExpressionProfile=NULL)
yAll<-
geneAnswersBuilder(AllEntrez,"org.Hs.eg.db",categoryType="KEGG",testType="hyper
G", pvalueT=0.1, FDR.correction=TRUE,geneExpressionProfile=NULL)
xxAll<-geneAnswersReadable(xAll)
yyAll<-geneAnswersReadable(yAll)

for (m in 1:length(TYPE_CHART)){
pdf(file=paste("AllSignifProbes_GO.BP_TOP5",TYPE_CHART[m],".pdf",sep=""),width=
14,height=9)
ChartPlotter(xxAll, chartType=TYPE_CHART[m])
dev.off()
print("Graph for KEGG")
pdf(file=paste("AllSignifProbes_KEGG_TOP5",TYPE_CHART[m],".pdf",sep=""),width=1
4,height=9)
ChartPlotter(yyAll, chartType=TYPE_CHART[m])
dev.off()
}
}
#End of the Function.....
```

ChartPlotter FUNCTION

```
ChartPlotter<-function (x, chartType = c("pieChart", "barPlot", "all"), sortBy
= c("geneNum","pvalue", "foldChange", "oddsRatio", "correctedPvalue"),
newWindow = FALSE, ...){
    chartType <- match.arg(chartType)
    sortBy <- match.arg(sortBy)
    if ((sortBy == "correctedPvalue") & !("fdr p value" %in%
        colnames(x@enrichmentInfo)))
        stop("input GeneAnswer instance does not contain corrected p value!!!")
    orderby <- switch(sortBy, geneNum = c("genes in Category",
        "TRUE"), pvalue = c("p value", "FALSE"), foldChange = c("fold of
overrepresents",
        "TRUE"), oddsRatio = c("odds ratio", "FALSE"), correctedPvalue = c("fdr
p value",
        "FALSE"))
    y <- x
    y@enrichmentInfo <- x@enrichmentInfo[order(x@enrichmentInfo[,
        orderby[1]], decreasing = as.logical(orderby[2])), ]
    y@genesInCategory <- x@genesInCategory[rownames(y@enrichmentInfo)]
    PlotCharts(y@enrichmentInfo, chartType = chartType, specifiedCols =
orderby[1],
        ylab = orderby[1], newWindow = FALSE, ...)
}
#End of the Function.....
```

PlotCharts FUNCTION

```
PlotCharts<-function (x, chartType = c("pieChart", "barPlot", "all"),
specifiedCols = c("genes in Category"),
    top = 5, newWindow = FALSE, ...)
{
    chartType <- match.arg(chartType)
    if (is.matrix(x) | is.data.frame(x)) {
        if (all(specifiedCols %in% colnames(x)) & (length(specifiedCols) ==
            1)) {
            if ((chartType == "pieChart") | (chartType == "all")) {
                if (newWindow)
                  x11()
                pie(as.numeric(x[1:top, specifiedCols[1]]), labels =
rownames(x)[1:top],
                  col = rainbow(top), radius = 0.9, main = paste("Top ",
                    top, " Categories Distribution based on ",
                    specifiedCols, sep = ""), ...)
            }
            if ((chartType == "barPlot") | (chartType == "all")) {
                if (newWindow | chartType == "all")
                  x11()
                barplot(as.numeric(x[1:top, specifiedCols[1]]),
                  names.arg = rownames(x)[1:top], space = 0.6,
                  col = rainbow(5), main = paste("Top ", top,
                    " Categories Distribution based on ", specifiedCols,
                    sep = ""), ...)
            }
        }
        else {
            stop("One or two given names are not column names in the original
matrix!")
        }
    }
    else {
        stop("The input data format is not a matrix or dataframe!")
    }
}
#End of the function.....
```

KG_HyperG FUNCTION

```
KG_HyperG<- function(GeneList, universe_GeneIds, annotation, categoryName,
pValcutoff, testDir){

#GeneList:   list of the genes to be tested. probe ids required.
#annotation: A string giving the name of the annotation data
#                  package for the chip used to generate the data.
#                  e.g. 'hgu133plus2'
#categoryName:      string giving the name of the category for testing.
#                  e.g. "KEGG", "PFAM"
#pValcutoff: p-value to be used for significance treshold.
#                  default is 0.05
#testDir:    direction of the test, "over" for overrepresentation and
#            "under" for underrepresentation.

annotation<-
match.arg(annotation,c("hgu133plus2.db","hgu133a.db","hgu133b.db","hgu95a.db","
hgu95b.db"))
categoryName<-match.arg(categoryName,c("KEGG","PFAM"))
testDir<-match.arg(testDir,c("over","under"))

if(is.vector(universe_GeneIds)){
universe_GeneIds<-universe_GeneIds
}
else{
return<-universe_GeneIds
}

assign("annotation", annotation)

require(annotation,character.only=TRUE)
require(KEGG.db)

switch (annotation,
hgu133plus2.db= assign("GeneList", hgu133plus2ENTREZID[GeneList]),
hgu133a.db= assign("GeneList", hgu133aENTREZID[GeneList]),
hgu133b.db= assign("GeneList", hgu133bENTREZID[GeneList]),
hgu95a.db= assign("GeneList", hgu95aENTREZID[GeneList]),
hgu95b.db= assign("GeneList", hgu95bENTREZID[GeneList])
)


GeneList<-as.vector(unlist(as.list(GeneList)))
GeneList<-GeneList[!is.na(GeneList)]
GeneList<-GeneList[!duplicated(GeneList)]


HypObj<-
new("KEGGHyperGParams",geneIds=GeneList,universeGeneIds=universe_GeneIds,
annotation=annotation,categoryName=categoryName,pvalueCutoff=pValcutoff,testDir
ection=testDir)
TestResult<-hyperGTest(HypObj)

results<-as.data.frame(TestResult@pvalues)

SignifKEGGs_ID<-rownames(results)[which(results<=pValcutoff)]
assign("SignifKEGGIDs", SignifKEGGs_ID, envir=.GlobalEnv)

rownames(results)<-KEGGPATHID2NAME[rownames(results)]

SignifKEGGsNAMES<-rownames(results)[which(results<=pValcutoff)]
```

```
ExpTitle<-paste("KEGG Pathways overrepresented in your data
:","**********","**********", sep="\n")
write(ExpTitle, file="KEGG_HyperGeometricTest_Results.txt")
write(SignifKEGGsNAMES, file="KEGG_HyperGeometricTest_Results.txt",
append=TRUE)


}
#End of the function...
```

GO_HyperG FUNCTION

```
GO_HyperG<- function(ontology, conditional, GeneList, universe_GeneIds,
annotation, categorySubsetIds, pValcutoff, testDir){



#ontology:          A string specifying the GO ontology to use. Must be
#                   one of "BP", "CC", or "MF".
#conditional: A logical indicating whether the calculation should
#                   condition on the GO structure.'TRUE' or 'FALSE'
#GeneList:          list of the genes to be tested. probe ids required.
#universeGeneIds: Object of class "ANY": A vector of gene ids in the
#                   same format as geneIds defining a subset of the gene
#                   ids on the chip that will be used as the universe for
#                   the hypergeometric calculation.
#                   If this is NULL or has length zero, then all gene ids
#                   on the chip will be used.
#annotation:  A string giving the name of the annotation data
#                   package for  the chip used to generate the data.
#                         e.g. "hgu133plus2.db"
#categorySubsetIds: Object of class "ANY": If the test method
#                         supports it, can be used to specify a subset of
#                         category ids to include in the test instead of
#                         all possible category ids.
#pValcutoff:  p-value to be used for significance treshold.
#                   default is 0.05
#testDir:           direction of the test, "over" for overrepresentation
#                   and "under" for underrepresentation.


library(Category)
library(GOstats)

annotation<-
match.arg(annotation,c("hgu133plus2.db","hgu133a.db","hgu133b.db","hgu95a.db","
hgu95av.db","hgu95b.db"))
categoryName<-c("GO")
datPkg<-DatPkgFactory(annotation)
testDir<-match.arg(testDir,c("over","under"))
ontology<-match.arg(ontology,c("BP", "CC", "MF"))


if(is.vector(universe_GeneIds)){
universe_GeneIds<-universe_GeneIds
}
else{
return<-universe_GeneIds
}


assign("annotation", annotation)

require(annotation,character.only=TRUE)


switch (annotation,
hgu133plus2.db= assign("GeneList", hgu133plus2ENTREZID[GeneList]),
hgu133a.db= assign("GeneList", hgu133aENTREZID[GeneList]),
hgu133b.db= assign("GeneList", hgu133bENTREZID[GeneList]),
hgu95a.db= assign("GeneList", hgu95aENTREZID[GeneList]),
hgu95b.db= assign("GeneList", hgu95bENTREZID[GeneList])
```

```
)


GeneList<-as.vector(unlist(as.list(GeneList)))
GeneList<-GeneList[!is.na(GeneList)]
GeneList<-GeneList[!duplicated(GeneList)]


HypObj<-new("GOHyperGParams",ontology=ontology, conditional=conditional,
geneIds=GeneList,universeGeneIds=universe_GeneIds,
annotation=annotation,categorySubsetIds=categorySubsetIds,categoryName=category
Name,datPkg=datPkg, pvalueCutoff=pValcutoff,testDirection=testDir)
TestResult<-hyperGTest(HypObj)

summGO<-summary(TestResult)
write.table(summGO, file="GO_hyperGTest_Results.xls", sep="\t")

}
#End of the function...
```

Grapher FUNCTION

```
Grapher<-function(ndata, foldchangeData, SignifKEGGIDs){

library(RColorBrewer)
library(org.Hs.eg.db)
library(RBGL)
library(grid)
library(KEGG.db)
library(KEGGgraph)

EntrezIDMaP <- hgu133plus2ENTREZID
TempNamesfc<-unlist(as.list(EntrezIDMaP[rownames(fc)]))
rownames(fc)<-TempNamesfc
fc<-as.data.frame(fc)
fc<-fc[which(!is.na(rownames(fc))),]

deKID <- translateGeneID2KEGGID(rownames(fc))


TempNamesndata<-unlist(as.list(EntrezIDMaP[rownames(ndata)]))
Tempndata<-ndata
rownames(Tempndata)<-TempNamesndata
ndata_framed<-as.data.frame(Tempndata)
ndata_framed<-ndata_framed[which(!is.na(rownames(ndata_framed))),]

allKID <- translateGeneID2KEGGID(rownames(ndata_framed))


FileList<-list()
for (i in 1:length(SignifKEGGIDs)){
FileList[i]<-
system.file(paste("extdata/hsa",SignifKEGGIDs[i],".xml",sep=""),package="KEGGgr
aph")
}
FileList<-FileList[is.na(charmatch(FileList,""))]
FileList<-unlist(FileList)



for (i in 1:length(FileList)){
       gg<-parseKGML(FileList[i])
       g<-KEGGpathway2Graph(gg,expandGenes=TRUE)
       gGeneID<-translateKEGGID2GeneID(nodes(g))
       gSymbol<-sapply(gGeneID, function(x) mget(x, org.Hs.egSYMBOL,
ifnotfound=NA)[[1]])

       for(j in 1:ncol(fc)){
              ar <- 20
              cols <- rev(colorRampPalette(brewer.pal(6, "RdBu"))(ar))
              logfcs<-fc[,j][match(nodes(g),deKID)]
              names(logfcs) <- nodes(g)
              logfcs[is.na(logfcs)] <- 0
              if(min(logfc)<=0){
              incol <- round((logfcs+(-min(logfcs)+1))*5)
              }
              else{
              incol<- round((logfcs)*5)
              }
              incol[incol>ar] <- ar
              undetected <- !nodes(g) %in% allKID
              logcol <- cols[incol]
```

```
            logcol[logfcs==0] <- "darkgrey"
            logcol[undetected] <- "yellow"
            names(logcol) <- names(logfcs)
            nA <- makeNodeAttrs(g, fillcolor=logcol, label=gSymbol, width=10,
height=1.2)
            par(mar=c(3,5,0,5), mgp=c(0,0,0))
            layout(mat=matrix(c(rep(1,8),2), ncol=1, byrow=TRUE))
#jpeg(file=paste(colnames(fc)[j],gg@pathwayInfo@title, ".jpeg", sep="_"))
            plot(g, "dot", nodeAttrs=nA,
main=paste(colnames(fc)[j],gg@pathwayInfo@title, sep=" "))
            image(as.matrix(seq(1,ar)), col=cols, yaxt="n", xaxt="n")
            mtext("down-regulation", side=1,  at=0, line=1)
            mtext("up-regulation", side=1,  at=1, line=1)
#dev.off()

      }

}
}
#end of the function...
```