UNIVERSITY OF GOTHENBURG

DEPARTMENT OF MEDICAL RADIATION SCIENCES SAHLGRENSKA ACADEMY

Age-Related Late Effects in Rat Thyroids After Internal ^{131}I Exposure

Ludwig Parling

Essay/Thesis:	30 hp
Program and/or course:	Sjukhusfysikerprogrammet – Självständigt arbete i radiofysik
Level:	Second Cycle
Term/year:	Spring/2023
Supervisor:	Eva Forssell Aronsson, Klara Insulander Björk, Anja Schroff
Examiner:	Magnus Båth
Keywords:	Thyroid, ^{131}I, late effects, internal exposure, age relation

Abstract

The Chernobyl nuclear accident left many questions in its wake. One of these questions was why children affected by the radioactive fallout seemed to have a higher incidence of thyroid cancer compared with adults in the same area. The purpose of this study was to compare proteomic changes in internally ^{131}I irradiated rats of two age groups, one young and one adult. This was performed through analysis of fresh frozen thyroid tissue from a previous animal study where young and adult rats were exposed internally to various amounts of ^{131}I. Analysis of the rats' thyroid proteome was achieved through liquid chromatography tandem mass spectrometry and analysed with Welch's t-test with Benjamini-Hochberg correction. Protein functional analysis and gene ontology (GO) term analysis were done with UniProt. Large variations in mainly the control group's data made comparison at the group level difficult. One protein's relative abundance, TRXR3, was found significantly changed in the young population at 0.5 kBq. TRXR3 is the rat equivalent to the TRXR3 found in humans and is a part of cellular oxidant detoxification and redox homeostasis. This would imply that oxidative stress had taken place, but the lack of other markers makes it difficult to draw a solid conclusion. Due to the lacking appearance at other doses, TRXR3 was not considered a good candidate biomarker for ^{131}I radiation exposure or absorbed dose.

Sammanfattning

Tjernobylolyckan skapade ett antal frågor, en av vilka var varför barn som påverkades av det radioaktiva nedfallet hade en högre incidens av sköldkörtelcancer jämfört med vuxna som verkade vara opåverkade av samma nedfall. Frågan är fortfarande utan svar. Syftet med studien var att undersöka proteomiska skillnader mellan vuxna och unga råttor som internt bestrålats med ^{131}I. Detta utfördes genom att analysera frusen sköldkörtelvävnad från en tidigare djurstudie där unga och vuxna råttor exponerades för olika mängder ^{131}I. Proteomiska förändringar mättes genom vätskekromatografitandemmasspektrometri och utvärderades via Welchs t-test med Benjamini-Hochberg metoden. Uniprot användes för identifiering av proteiner. Eftersom datan innehöll stora variationer var det svårt att dra slutsatser. TRXR3 var det enda upphittade proteinet och befann i gruppen unga råttor, 0.01 Gy. Proteinet är väldigt lik det mänskliga proteinet TRXR3. Den tar del i bland annat cellulär oxidationsavgiftning och redox homeostas vilket skulle implicera oxidativ stress men bristen på andra markörer för det gör slutsatser svåra att dra. Eftersom TRXR3 bara upphittades vid en dos ansågs den inte vara en kandidat för att biomarkera ^{131}I exponering eller absorberad dos.

Table of contents

Introduction

After the Chernobyl nuclear accident, the cancer rates of the thyroid, most likely caused by internal exposure to ^{131}I, were observed to dramatically increase among children below 14 years of age at the time of the accident (UNSCEAR, 2017). At the same time, the thyroid cancer incidence in individuals considered to be adults at the time of the accident remained seemingly unaffected by the radioiodine (UNSCEAR, 2017). While research has been conducted on identifying the radiobiological response after internal low dose exposure to ^{131}I (Larsson, et al., 2020; Larsson, et al., 2022), the mechanisms which make children more susceptible to ionizing-radiation-induced thyroid cancer remain mostly unknown (UNSCEAR, 2017).

Background

The Iodine Isotope ^{131}I

^{131}I is an isotope with atomic number 53 and is used to treat certain thyroid-related illnesses, e.g., cancer and hyperthyroidism (Berg, G., et al., 2007). It emits both gamma and beta radiation, and decays into ^{131}Xe, a stable element. Since ^{131}I decays by beta emission with a mean energy of ca. 190 keV for the beta particles and ca. 364 keV for the gamma photons and has a physical half-life of about 8 days, it fits quite well into the field of theranostics. ^{131}I is suitable for regional use against tumors due to the short range of the beta particles and functionally useful half-life. The gamma emission expands the use of ^{131}I making imaging of the biokinetics and dose estimations possible, giving it the capability to be a theranostic tool.

A special use for ^{131}I is the treatment of thyroid related issues, such as thyroid cancer or some types of overt hyperthyroidism. Due to the thyroid requiring iodine to produce the hormones triiodothyronine (T3) and thyroxine (T4), the thyroid will naturally accumulate iodine (Berg, G., et al, 2007). This makes ^{131}I cheap to use and efficient at treating thyroid related disorders as no carrier needs to be used to bring the radiopharmaceutical to the thyroid and no chelator is needed, minimizing chemical toxicity.

Radiation Protection

Protection against unwanted radiation exposure, not only from ^{131}I, usually comes down to three factors, exposure time, range to the source, and shielding against the radiated particles. Reducing exposure time and increasing range to the source are simple ways to reduce absorbed dose and risk for adverse effects as simply moving away from the affected area will do in applicable cases. When this is not possible, however, shielding takes a greater role in reducing the risk of aftereffects. In the context of ^{131}I, a common protective measure are potassium iodine tablets. They work on the premise that the thyroid can be saturated with iodine so that additional iodine will not be retained in the thyroid. This would protect from retaining radioactive iodine in the thyroid and lessen exposure time if it cannot be avoided.

When exposure cannot be prevented, biomarkers can be a helpful tool in showing e.g., exposure, dose, and the presence of tumors. The specific information they offer could infer a better understanding of biological effects and risks at play after exposure to ionizing radiation. Several studies have proposed candidate biomarkers for internal exposure to ^{131}I in children, such as, parvalbumin (PVALB) and thyroglobulin (TG) (Larsson, et al., 2022; UNSCEAR, 2017). While there are plenty of studies showing significant responses in their suggested biomarkers, most of them show inconsistencies in which candidates are suggested (Larsson, et al., 2020; Larsson, et al., 2022; UNSCEAR, 2017).

The Thyroid

The thyroid is an organ which sits around the top, and in front, of the trachea. It weighs about 20 g in humans and consists of two lobes which together form a butterfly shape (Berg, G., et al., 2007). It is an endocrine organ which mainly regulates growth, metabolic rate (short and long-term) and protein synthesis using the hormones T3, T4, and calcitonin (Berg, G., et al., 2007). This can be done through directly affecting proteins and hormones, or by affecting gene expression (Berg, G., et al., 2007). A main component in T3 and T4 is iodine (Berg, G., et al., 2007).

The thyroid consists of follicular and parafollicular cells, and follicles (Berg, G., et al., 2007). Parafollicular cells, also known as C-cells, are scattered throughout the thyroid and secrete calcitonin when stimulated (Berg, G., et al., 2007). Follicular cells enclose the follicles which produce the hormones T3 and T4 when stimulated by thyroid-stimulating hormones (TSH) (Berg, G., et al., 2007).

The thyroid works as a deposit for iodine (Berg, G., et al., 2007). About 30% of the consumed iodine is transported to the thyroid while the rest is excreted through urine (Berg, G., et al., 2007). This feature has been heavily relied upon when treating the thyroid using stable iodine, as well as internal irradiation by ^{131}I (Berg, G., et al., 2007). However, 30% is not necessarily what the thyroid will absorb. The supply of iodine to the thyroid would eventually lead to saturation, as such the thyroidal intake of iodine would decrease. This would suggest a relation between thyroid iodine saturation and thyroidal uptake of iodine.

Hyper-/Hypothyroidism

There are several implications ionizing radiation can have in the thyroid. A very common late effect, in the context of therapy, is hypothyroidism. Hypothyroidism is an umbrella term for various states which cause the thyroid to produce less hormones than expected. Be it external irradiation or internal irradiation with ^{131}I this is a frequent phenomenon, although the absorbed dose tends to be in the therapeutic range for most cases (Hancock, et al., 1995). The most common time of appearance, 2-3 years after exposure, may make some hesitate to call it a late effect. However, the time range in which it may occur is between a few months to over 20 years, in some cases, with about half of the cases occurring after 5 years (Colevas, et al., 2001; Illés, et. al., 2003; Mercado, et al., 2001). The frequency of occurrence of hypothyroidism increases with an increased dose (Metso, et al., 2004; Vogelius, et al., 2011).

Hyperthyroidism may also occur, while quite a bit less frequent than hypothyroidism (Hancock, et al., 1991; Fleming, et al., 1985). Hyperthyroidism is an umbrella term used to classify states of overproduction of thyroidal hormones. Given external radiation therapy to the neck with therapeutic doses, hyperthyroidism shows itself at similar time intervals as hypothyroidism (Hancock, et al., 1991) while it shows itself months after internal ^{131}I radiotherapy (Dunkelmann, et al., 2004; Nygaard, et al., 1997). It is currently unclear why this is (Nagayama, et al., 2018). There is also evidence which supports a correlation between dose and incidence of hyperthyroidism for high doses (Dunkelmann, et al., 2004; Nygaard, et. al., 1997), but this is disputable for low to moderate doses (Imaizumi, et al., 2017).

Thyroid Cancer

In 2022 the number of thyroid cancer cases, in the US, was estimated at 2.3% of all cancer cases (NCI, 2023). In Sweden, in the year 2021, the number of new thyroid cancer cases was reported to be 19.94 persons per 100 000 (Cancerregistret, 2022). Thyroid cancer is also more common among women, compared with men, where women usually represent two-thirds to three-quarters of all cases and is most frequent in people in the age range of 30-74 (Coleman, et.al, 1993; NCI, 2023). The incidence trend for thyroid cancer around the world is a mix of decrease and increase in different countries while the general trend is pointing to an increase (Coleman, et.al, 1993; NCI, 2023; Cancerregistret, 2022). In the US the incidence of thyroid cancer was

observed at 7.5 persons per 100 000 in the year 2000 and 13.8 persons per 100 000 in 2019 (NCI, 2023). For the period, 2000 to 2019, the measured difference was an increase of 10.77 persons per 100 000, in Sweden (Cancerregistret, 2022). The mortality, however, has been decreasing. In the US, between the years 2000 and 2015, the 5-year relative survival climbed by 2%-units from an observed 96.6% to an observed 98.6% (NCI, 2023). This is mostly attributable to improvements in diagnostic imaging techniques which allow for imaging non-palpable nodules (Cabanillas, et al., 2016).

Thyroid cancer can be divided into neuroendocrine C-cell-derived and follicular-cell-derived thyroid cancer (Cabanillas, et. al., 2016). Follicular-cell-derived thyroid cancer can be divided further into differentiated and undifferentiated thyroid cancer. Yet another subdivision can be created by dividing differentiated thyroid cancer into papillary, follicular, and Hurthle cell thyroid cancer (Cabanillas et al., 2016). The most common type of thyroid cancer is differentiated thyroid cancer, representing more than 95% of cases, while the most common type in the subdivision is papillary thyroid cancer (Howlader, et.al, 2016; Cabanillas, et.al, 2016). Fortunately, it is the papillary variant that is the most benign while the other types of differentiated thyroid cancer are high risk and are known to metastasize and spread into lung and bone tissue (Cabanillas, et.al, 2016).

In the case of childhood thyroid cancer caused by ionizing radiation, the most common variant is papillary thyroid cancer (Bresciani, et.al, 2019) as this is the most common thyroid cancer form induced by ionizing radiation, which was also observed in children after the Chernobyl nuclear powerplant accident (Thomas and Yamashita, 2017). More specific details on the carcinogeneic development, like differences between spontaneous and ionizing-radiation-induced thyroid cancer, remain unknown (UNSCEAR, 2017).

Sprague-Dawley Rats

The Sprague-Dawley rats are a common and well-documented species of rats commonly used in animal experimentation. They tend to gestate in 21 days after which they enter a period of sexual maturation until week 10 after birth consisting of the phases, neonatal, infantile, juvenile, peripubertal, and adolescents, lasting from day 0-7, 7-21, 21-35, 35-55, and 55-70 respectively (Ghasemi, et.al, 2021; Beckman, 2003; Bell, 2018; Marty, et.al, 2003; Picut, et.al, 2018; Semple, et.al, 2013; Vidal, 2017). The rats tend to enter adulthood at the end of sexual maturation at about week 10 starting with the emerging adulthood and continuing into young, middle, older, and late adulthood at about the ages of 70, 150, 300, 600, and 730 days respectively (Ghasemi, et.al, 2021; Quinn, 2005; Stanley and Shetty, 2004).

Proteomic Analysis

DNA is often referred to as a sort of blueprint, of which proteins are the product. As such, there is a profound connection between genes and proteins, but variations in expression between the two are frequent as the regulators mediating the transcriptional and translational processes are plentiful (Maier, et.al, 2009; de Sousa Abreu, et.al, 2009). It is therefore difficult and not advised to draw conclusions on gene expression from proteomic analyses. The proteome can, however, still be used to observe the consequences of gene expression change. Other studies have used proteomics to investigate the mechanics of late effects after exposure to ^{131}I (Larsson, et.al, 2020; Larsson, et.al, 2022). Their results showed statistically significantly changed expression in several proteins, but few consistencies in which proteins were significant and changed in the same manner. This can be used as a reference for what is to be expected from proteomic studies on the thyroid after internal exposure to ionizing radiation from ^{131}I. In this study the effect of ionizing radiation from ^{131}I on the thyroid was investigated through proteomic analysis.

A method for protein analysis is mass spectrometry (MS). MS works by deviating charged particles in an electromagnetic field. The particle's trajectory is determined by the particle's inertia and charge. The electromagnetic field produces a force on the charged particle which deviates the path while the inertia resists the change in momentum giving the particle a distinct path. This gives information on the particle's mass-to-charge ratio (m/z) which aids in the identification of the particle. There are, however, times where the m/z ratio is similar enough between particles that a distinction cannot be made. Two charged compounds can have different masses and different charges while still having the same m/z ratio. While it is difficult to completely remove the problem, increased performance can be achieved by coupling two or more MS analytical units together and fragmenting the particles into smaller constituents. This is called tandem MS (MS/MS). For substances still containing the possibility of indistinguishable MS profiles, liquid chromatography (LC) can help with this problem by providing refining and sorting information. In short, LC operates by running a mobile phase, in which the analyte is dissolved, through a stationary phase. The contents of the mobile phase will be separated by some attribute, like size, electric charge, solubility, etc., by interaction with the stationary phase and the product (eluent) will exit the stationary phase according to the sorting order. By using LC to discriminate substances in a solution through separation, the desired substance can be extracted, enabling analysis through MS while also contributing extra information in the form of retention time (RT) (Kailasam, 2021). Analysis of proteins can be done by cleaving them into smaller peptides and then attaching so called tandem mass tags for the simultaneous analysis of multiple samples, also known as multiplexing (Zhang and Elias, 2017). Large proteins may have similar RT and m/z so directly analyzing them via LC-MS/MS is a difficult task. The cleavage of proteins into peptides can be archived by e.g., trypsin, cleaving at specific amino acid junctions. Peptides are more easily identifiable as the m/z value for peptides makes for fewer similar results than proteins. Another advantage is that bigger proteins are difficult to dissolve in LC-MS/MS solvents and can therefore not be analyzed while peptides can. The tandem mass tags fragment in a particular manner which creates reporter ions specific for each tag. The ions become visible in the mass spectrum and makes the peptides'

abundances differentiable (Zhang and Elias, 2017). Analysis results are in the form of abundance from both LC and MS analysis and are plotted against the RT and m/z values in what is known as a total ion chromatogram (TIC) (Kailasiam, 2021). The TIC identifies the parent compound by matching fragmentation footprints at specific RTs with the parent's expected fragmentation pattern and RT. Quantification is performed by taking the area under a relevant MS peak for the relevant RT (Kailasiam, 2021).

A previous study on analyzing the effects of ^{131}I in the thyroid by Larsson, et.al, (2020) would suggest that there is a statistically significant change in the relative abundance of several proteins. It was found that SORBS2, ACADL, TG, and TPO were significantly overexpressed at low doses. The GO term analysis, which shows the processes and functions related proteins, revealed an immune response suggested by the discovered changes. There was found to be a dose-related response correlating with some of these proteins, although not monotonic.

Purpose

The purpose of this study was to identify and quantify changes in the proteome and causal effects in thyroid glands of rats irradiated internally with ^{131}I at young and adult age to achieve low to medium absorbed doses, with regards to late effects. Late effects were considered proteomic change which occurred due to ionizing radiation after 12 months.

Method

Animal Experiment

The tissues used in this study were taken from a previous study by Larsson, et al. (2020). The model organism for observing protein abundance change in this study were male Sprague Dawley rats from Charles River Laboratories International in Salzfeld, Germany. Groups were created as seen in Table 1. Rats were divided based on their age into two groups, young and adult, containing 24 and 27 rats respectively. The adult rats were intra-venously injected at the age of 17 weeks while young rats were injected at 5 weeks. The two age groups were divided into four activity groups, 0.5, 5, 50, and 500 kBq, and one control group which received saline solution. ^{131}I was injected and the rats' thyroid dose was calculated and associated with their activity group. The absorbed dose to the thyroid was calculated through the MIRD formalism (Mattsson, et.al, 2015; Spetz, et.al, 2013). Euthanasia was performed 12 months after injection and the rats were dissected, thyroids collected and flash frozen. The rats had free access to food and water and new equipment was used for every thyroid. The thyroids were cut into three pieces in a fumehood using sterile equipment. One of the pieces was sent for proteomic analysis. The animal experiment performed in this study has been approved by the Ethical Committee on Animal Experiments in Gothenburg Sweden (Permit Number: 146-2015).

Table 1: Experiment design and group size (parenthesized).

	Activity (kBq)	***-***	***0.5***	***5***	***50***	***500***
Dose to thyroid	**Adult**	0 Gy (5)	0.007 Gy (5)	0.07 Gy (6)	0.7 Gy (5)	7 Gy (6)
	Young	0 Gy (6)	0.01 Gy (5)	0.1 Gy (4)	1 Gy (5)	10 Gy (4)

Sample Preparation

The samples used in this study were placed in sterile test tubes and sent to the Sahlgrenska Academy Proteomics Core Facility where they were homogenized and prepared for LC-MS/MS analysis. Homogenization was performed by ultrasonication with a Covaris ML230 in lysis buffer (50 mM triethylammonium bicarbonate, TEAB, 2% sodium dodecyl sulfate). Protein concentration was measured using Pierce BCA Protein Assay (Thermo Scientific) on a Benchmark Plus microplate reader (BIO-RAD Laboratories). 40 µg protein samples were reduced for 30 minutes at 56ºC in 10 mM dithiothreitol followed by 10 minutes of room temperature alkylation using 20 mM iodoacetamide and 10 minutes of quenching with 10 mM dithiothreitol. Protein samples were placed in washed hydrophilic and hydrophobic carboxylate modified Sera-Mag™ SpeedBeads from Cytiva with a beads-to-protein ratio of 10:1. Protein and peptide clean-up dictated the SP3 workflow for the manufacturer-provided MS protocol. Proteins were precipitated on the beads with 100% ethanol and washed with 80% ethanol followed by being dried at room temperature. Digestion was done by incubating overnight at 37ºC while shaking after adding 50 mM TEAB and 1.2 µg LysC with trypsin (Promega) after which an additional 2 µg of trypsin (Thermo Fisher Scientific) for digestion over 4 hours. TMTpro 18-plex isobaric mass tagging reagents from Thermo Fisher Scientific were used to label the supernatant in accordance with manufacturer instructions. A reference was created by pooling the samples of a plex. Peptide purification was performed according to manufacturer instructions with Pierce peptide desalt spin column and HiPPR detergent removal kit from Thermo Fisher Scientific. A Dionex Ultimate 3000 UPLC reverse-phase chromatograph from Thermo Fisher Scientific was used for TMT fractionation. Solvent A was 25% ammonia and solvent B was 84% acetonitrile. A reversed-phase XBridge BEH C18 column (3.5 µm, 3x250 mm, Waters Corp.) and a stepped solvent B gradient from 3% to 54% over 65 minutes followed by an increase to 80% solvent B with a 200 µl//min flow was used for peptide separation. LC-MS3 analysis was prepared by reconstituting evaporated 36 fractions, which were combined from 72 primary fractions, in 3% acetonitrile and 0.1% trifluoroacetic.

Sample Analysis

The fractions were analysed via LC-MS3 in an Orbitrap Fusion Lumos Tribrid MS with an Easy-nLC1200 LC system and FAIMS Pro ion mobility system from Thermo Fisher Scientific. A Thermo Fisher Scientific Acclaim Pepmap 100 C18 column (100 µm x 2 cm, 5 µm particle size) was used to trap the peptides. Separation was done on a Dr. Maisch Reprosil-Pur C18 analytical column (35 cm x 75 µm, 3 µm particle size) with a stepped gradient on the interval [4,80] % acetonitrile in 0.2% formic acid over 75 minutes at a flow of 300 nl/min. The same data-dependent settings were used for the FAIMS Pro when alternating between the compensation voltages, -50 and -70. The precursor mass spectra had a resolution of 120 000 and an m/z range of 375-1500. The most abundant precursor charges between 2 and 7 were isolated using a cycle time of 1.5 seconds and an m/z window of 0.7 and fragmented by

collision-induced dissociation at 30%. Recording of fragment spectra was done at Rapid scan rate. Further MS3 fractionation was done for the 10 most abundant MS2 fragments. The 10 fragments were isolated using multi notch isolation. High energy collision dissociation at 55% was used to perform MS3 fractionation. Recording resolution was 50 000 and m/z range was [100,500].

Identification and relative quantification were performed with Thermo Fisher Scientific's, Proteome Discoverer 2.4. SwissProt (8177 entries, April 2023) was used to match the data against *Rattus norvegicus*. Precursor tolerance and fragment ion tolerance was 5 ppm respectively 0.5 Da for a Sequest database matching. One missed cleavage was accepted for tryptic peptides. Methionine oxidation was put as a variable modification. Fixed modifications were cysteine carbamiodomethylation, TMTpro on lysine, and peptide N-termini. A percolator was used as a filter for PSM validation with a 1% FDR. TMT reporter ions were identified in the MS3 HCD spectra with 3 mmu mass tolerance and TMT reporter intensity for each sample were normalized on the total protein contents. SPS Mass Match threshold was 65%. Sequest XCorr threshold was 1.2. Unique peptides were used for relative quantification. Proteins were required to pass with an FDR of 5%.

Welch's t-test

A suitable statistical test for analysis of change in protein expression is Welch's t-test. The central limit theorem can be used to assume that the tissue protein abundance between tissue samples is approximately normally distributed, and a t-distribution can therefore be used to describe a limited sample size of tissue samples. Given the study of protein expression before vs. after some treatment, unequal variance between the two may also be assumed as there is no guarantee that the treatment retains the variance.

Welch's t-test, or two-sample t-test with unequal variance, is a statistical test used for comparing two populations where normal distribution can be approximated, but equal variance between the observed populations cannot. J.S. Milton (1995) states that the comparison is described using the statistic,

$$\overline{D} = \overline{X} - \overline{Y} \quad (1)$$

where $\overline{X}$ and $\overline{Y}$ are the approximately normally distributed variable for the two populations being compared. If $\overline{X}$ and $\overline{Y}$ consists of n_X respectively n_Y points, the resulting test statistic for hypothesis testing will be,

$$T = \frac{\overline{D} - \mu_d}{S} \quad (2)$$

where $S = \sqrt{\frac{S^2_X}{n_X} + \frac{S^2_Y}{n_Y}}$ is the standard deviation for $\overline{D}$, the mean for the sample difference, and $\mu_d = \mu_x - \mu_y$ is the hypothesis testing variable, which is usually set to zero (no change) for the null hypothesis (H_0). This distribution follows a $T_{(i-1)}$ distribution, granted H_0 is true.

Benjamini-Hochberg Procedure

Given some H_0 and an approximately normally distributed dataset, there will be a value given by Equation (2) which corresponds to the probability of the dataset occurring at random, given that H_0 is true. This probability is commonly known as the p-value. One implication of the existence of this probability is that some of the data will speak for accepting H_0, even if it is not true. When H_0 is not true, accepting it is known as committing a type I error.

Since sampling datapoints which speak for committing a type I error is impossible to avoid, using the data will undoubtedly influence statistical accuracy. In situations where statistical analyses compound one another, so does the errors and tracing the errors may indeed be difficult. A common way of controlling the type I error incidence is by utilizing the Benjamini-Hochberg (BH) method.

The BH method operates on the distribution of p-values among all possible p-values. It can be proven that when samples come from a single distribution, the p-values of the samples should be evenly distributed across all possible p-values, e.g., there should be equally many p-values between 0 and 0.05 as there are between 0.5 and 0.55. When samples are taken from two different distributions, the distribution of p-values is shifted to the lower end, i.e., there are more smaller p-values than there are larger.

As described by Agresti (2012), the BH procedure starts by ordering all p-values in ascending order and assigning a rank to each p-value in the same order, after which the BH critical value, P_{BH}, is calculated. This is performed by multiplying the chosen false discovery rate, q, with the rank, k, of the p-values and dividing by the total number of p-values, n.

$$P_{BH} = q\frac{k}{n} \quad (3)$$

The next step is to search for the largest p-value that is still smaller than the BH critical value. If this value is positioned such that no following p-value is smaller than their respective BH critical value, this is the cut-off point. All values above the cut-off are to have H_0 rejected with a q chance of committing a type I error.

Using these values, a graph which plots the p-value to the rank can be generated which either has an exponential or linear shape. If the shape is linear, there are equally many p-values for any given p-value bin and the data speaks for H_0. If the shape is exponential, there are more smaller p-values than there are larger and therefore data which speaks for rejecting H_0.

Statistical Analysis

The data was evaluated using Welch's t-test with a BH adjusted p-value in the program R 4.2.2. The data was loaded into Microsoft Excel, where reformatting was performed, and relevant data was sorted into a more interface friendly state for the user and R. The processed data was then loaded into R where a Welch's t-test was performed, whereby each group with an injected activity over 0 kBq in the adult and young categories had their protein abundance tested against the abundance of the 0 kBq group in their respective category. The p-values were saved and the BH correction was applied. The false discovery rate was set to 5% and all BH corrected p-values below this limit was considered statistically significant. A multiplicative change factor of protein abundance relative the reference, known as fold change (FC), limit of 1.5 was set to indicate that relevant changes in protein abundance for a protein was considered an increase or decrease of 50%.

Proteins were identified using UniProt (EMBL-EBI, 2023). The proteins with a statistically changed abundance and FC of 1.5 or more had their accession codes fed into UniProt where information on GO annotations were collected. The same was done for the human equivalent of the rat proteins by identifying the gene which produces a specific protein and checking if it is present in humans and what protein the gene produces in humans. Tissue expression data for the human equivalent proteins were collected from Human Protein Atlas (2023; Uhlén, et.al, 2015).

Results

The results were a list of each rat's protein abundance, measured in FC, for every detectable protein. 7386 proteins were detected, out of which 4681 were analyzed with statistics and the rest were omitted due to insufficiently accurate quantity data. Every group of rats had a mean FC and standard deviation for the occurrence of every analyzed protein. The irradiated groups also had BH corrected p-values and mean values which were relative to the groups' control group.

Study Results

The following images of the results are volcano plots of the proteins' BH corrected p-values, $-log_{10}(P)$, plotted against the groups' mean abundance relative to the control mean abundance of the same protein, $-log_2(FC)$, at the specified dose. The points in the volcano plot represent various proteins' BH corrected p-value and FC. Vertical lines mark the FC limit of 1.5 and horizontal lines mark the FDR limit of 5%. Grey dots represent proteins without a statistically significant change in abundance and not enough of a difference in abundance to be considered a relevant change from the mean of the control groups. Green dots represent a sufficient difference in protein abundance from the mean of the control groups without being statistically significant. Red dots represent proteins that are statistically significant as well as differing sufficiently in abundance from the mean of the control groups.

Adult Rats

Figure 1 shows volcano plots of the BH corrected p-values against the mean relative FC for the analyzed proteins in each activity group for the adult groups. The sub-figures A, B, C, and D show results for the activity groups 500, 50, 5, and 0.5 kBq respectively.

As seen in Figure 1, no significant proteomic change was registered for the adult rats at the tested doses. Plenty of large fold changes were observed, but nothing was measured as significant with a 5% FDR and an FC cut-off of 1.5. The figure also shows that there was no significant change at any FC cut-off, even for a higher FDR. The FDR seemed to also be increasing for an increased dose, since the $-log_{10}(p)$ generally seemed to decrease with an increased dose, creeping extremely close to 0 at 50 and 500 kBq.

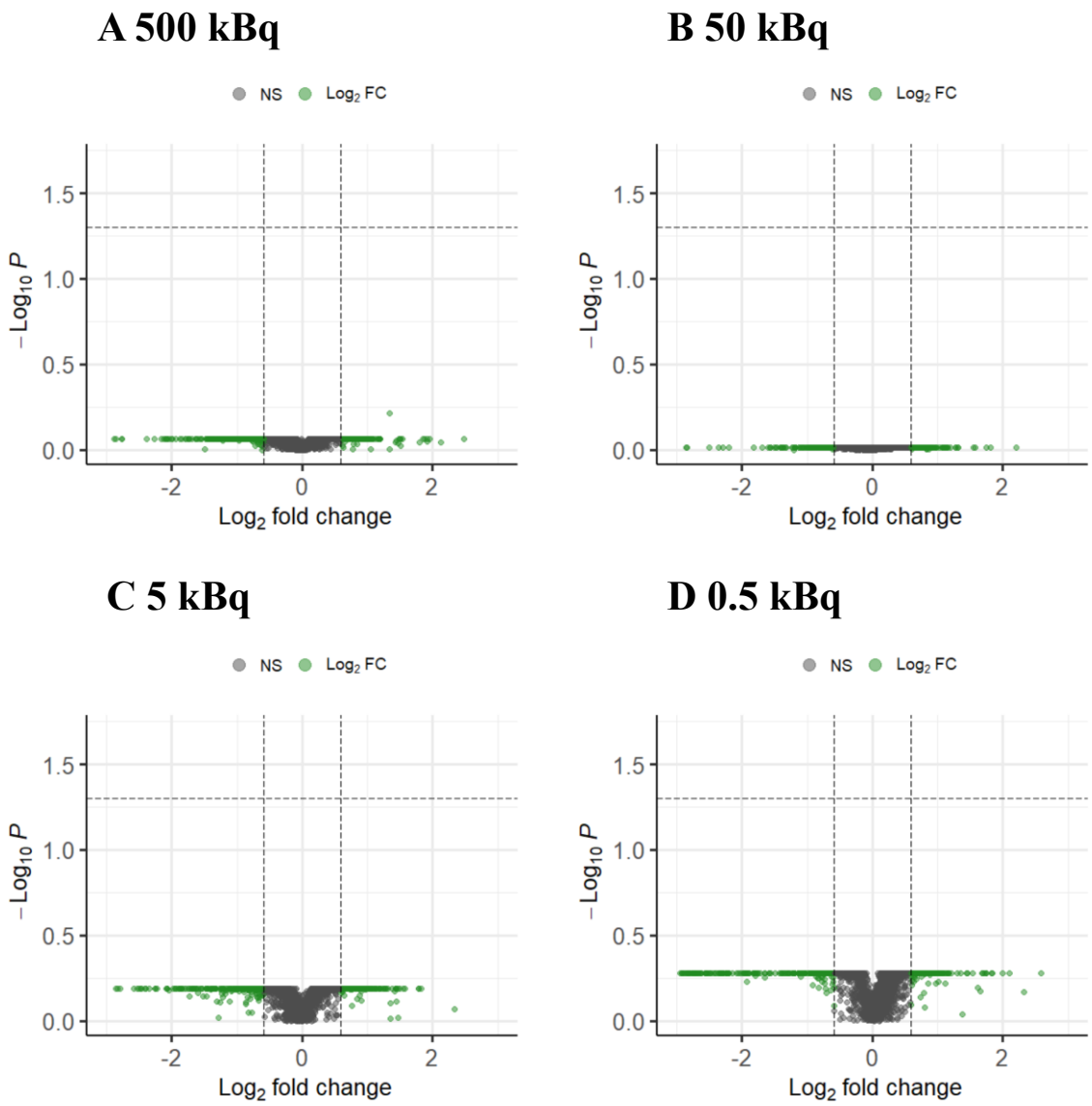

Figure 1: Volcano plots for the tested proteins for the adult rats, for the different activity groups 0.5, 5, 50 and 500 kBq. NS stands for non-significant, Log_2 FC stands for relevantly changed FC. The x-axis shows the log_2(fold change) and the y-axis shows the -log_{10}(B.H. corrected p-value).

Young Rats

Figure 2 shows volcano plots of the BH corrected p-values against the mean relative FC for the analyzed proteins in each activity group for the young main group. The sub-figures A, B, C, and D show results for the activity groups 500, 50, 5, and 0.5 kBq respectively.

Like for the adult rats, the FC varied, but only one protein was discovered with a 5% FDR and a FC cut-off of 1.5 (Figure 2). The *-log_{10}(p)* seemed to decrease with increased dose and reach close to 0 at 1 and 500 kBq. The only statistically significantly altered abundance registered was an upregulation in a protein called thioredoxin-disulfide reductase (TRXR3). This was only seen for 0.5 kBq and was not present at any other dose.

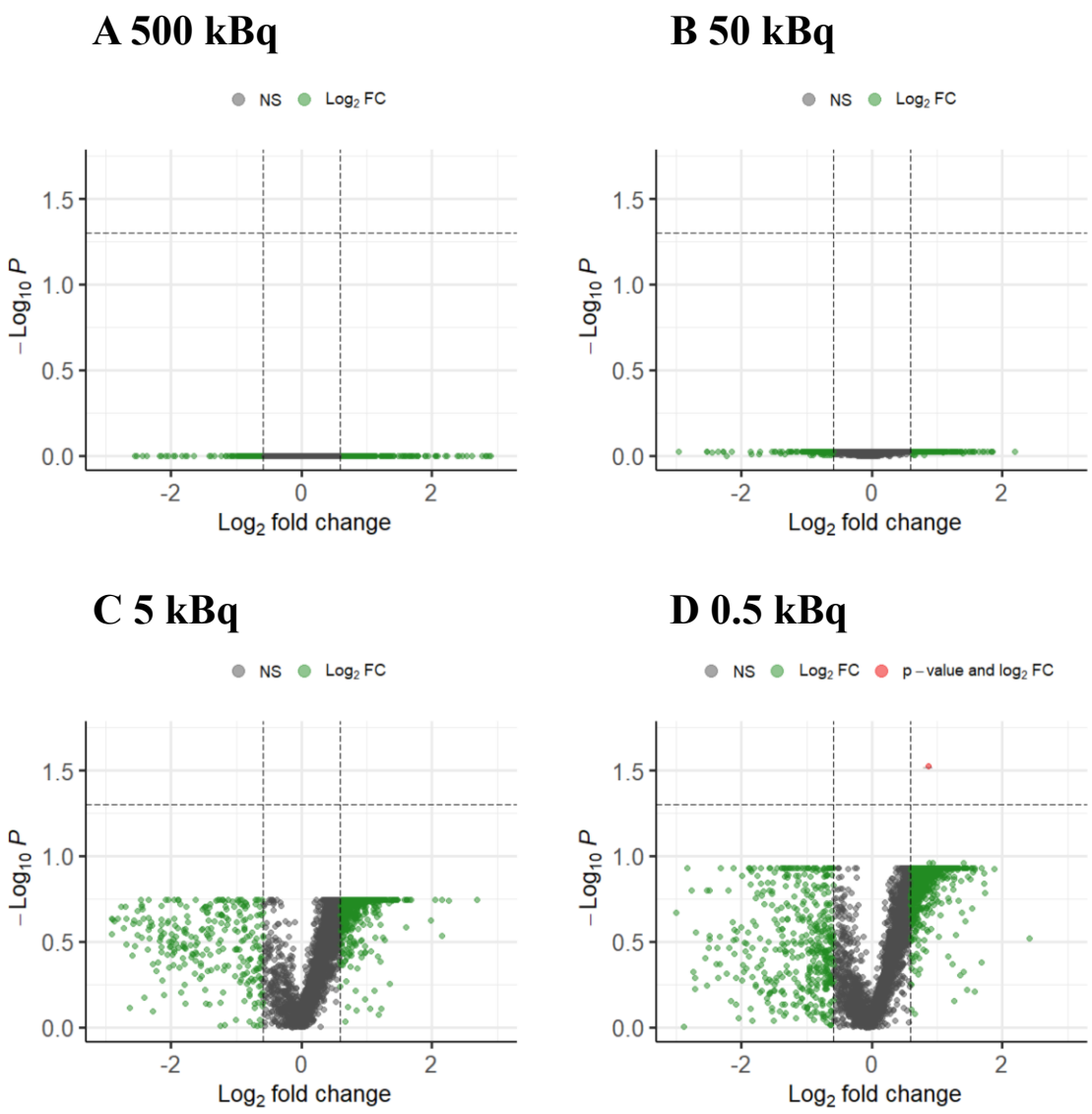

Figure 2: Volcano plots for the tested proteins for the young rats, for the different activity groups 0.5, 5, 50 and 500 kBq. NS stands for non-significant, Log_2 FC stands for relevantly changed FC. The x-axis shows the log_2(fold change) and the y-axis shows the -log_{10}(B.H. corrected p-value)

Result Accuracy

The following figures are plots of the p-values for each protein against their respective rank as assigned by the BH method. These plots contain information on accuracy via the shape of the trend in p-value to rank.

Adult Rats

Figure 3 shows the p-value-to rank for the sub-figures A, B, C, and D which show results for the activity groups 0.5, 5, 50, and 500 kBq respectively. Overall, the curves have an S-like shape (sigmoidal), starting slow and then increasing more rapidly. This trend is becoming less and less prominent with higher activity, whereby the last curve looks almost linear. The curves have a gentle upward slope at first, indicating more proteins for lower p-values. As the rank continues to rise, the slope steepens, showing fewer proteins at high p-values.

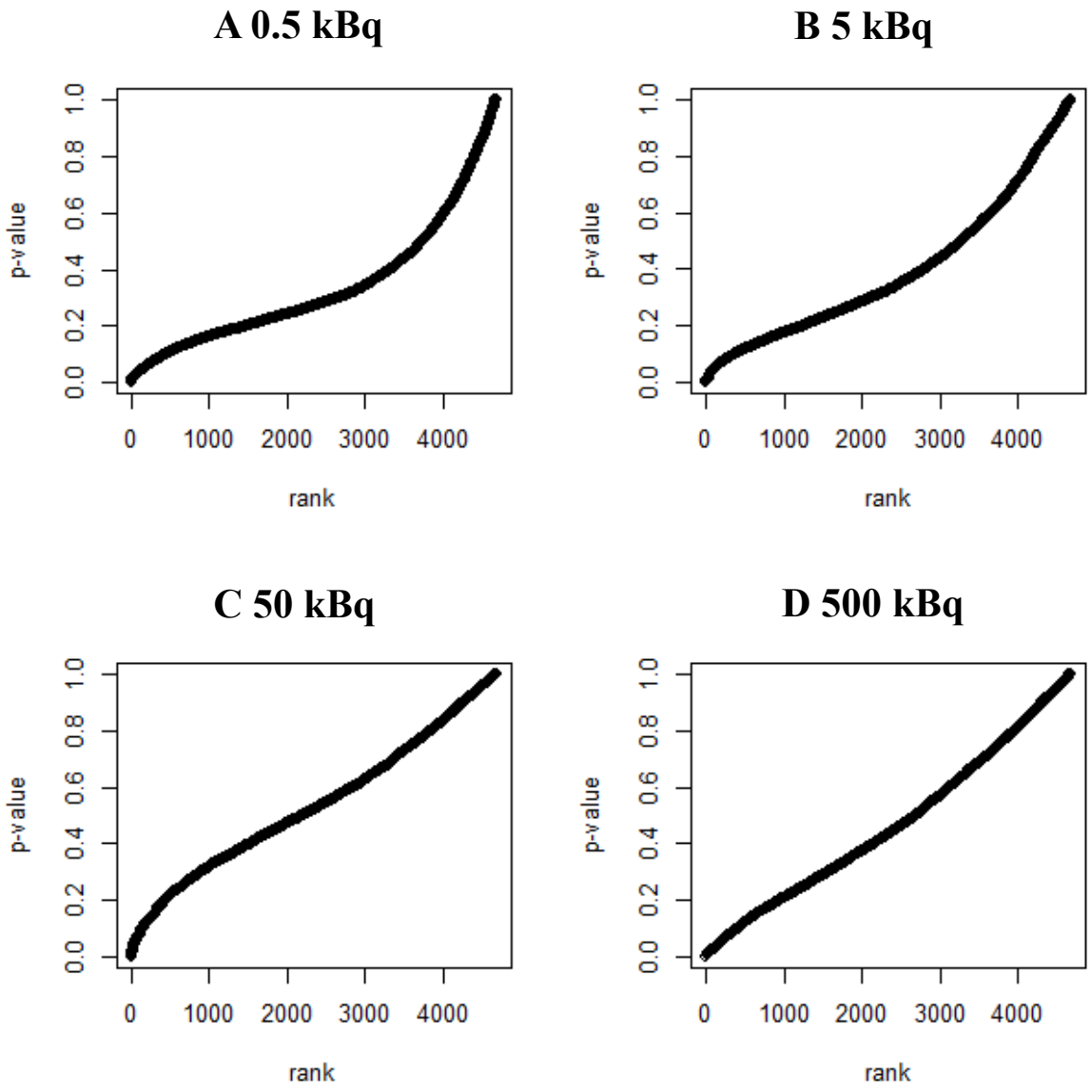

Figure 3: P-values ordered in ascending order plotted against their rank for the adult rats at the different investigated activity levels 0.5, 5, 50 and 500 kBq.

Young Rats

Figure 4 shows the p-value-to rank for the sub-figures A, B, C, and D which show results for the activity groups 500, 50, 5, and 0.5 kBq respectively. Subfigures A and B look similar to an exponential curve, with more proteins at lower p-values. As the rank increases, the slope becomes steeper, indicating a faster increase in p-values. The opposite is true for subfigures C and D, with more proteins at lower p-values, followed by a progressively increasing amount of proteins at higher p-values showing itself as a flattening trend. This is particularly pronounced for subfigure D, and the trend resembles a logarithmic curve.

.

Variance

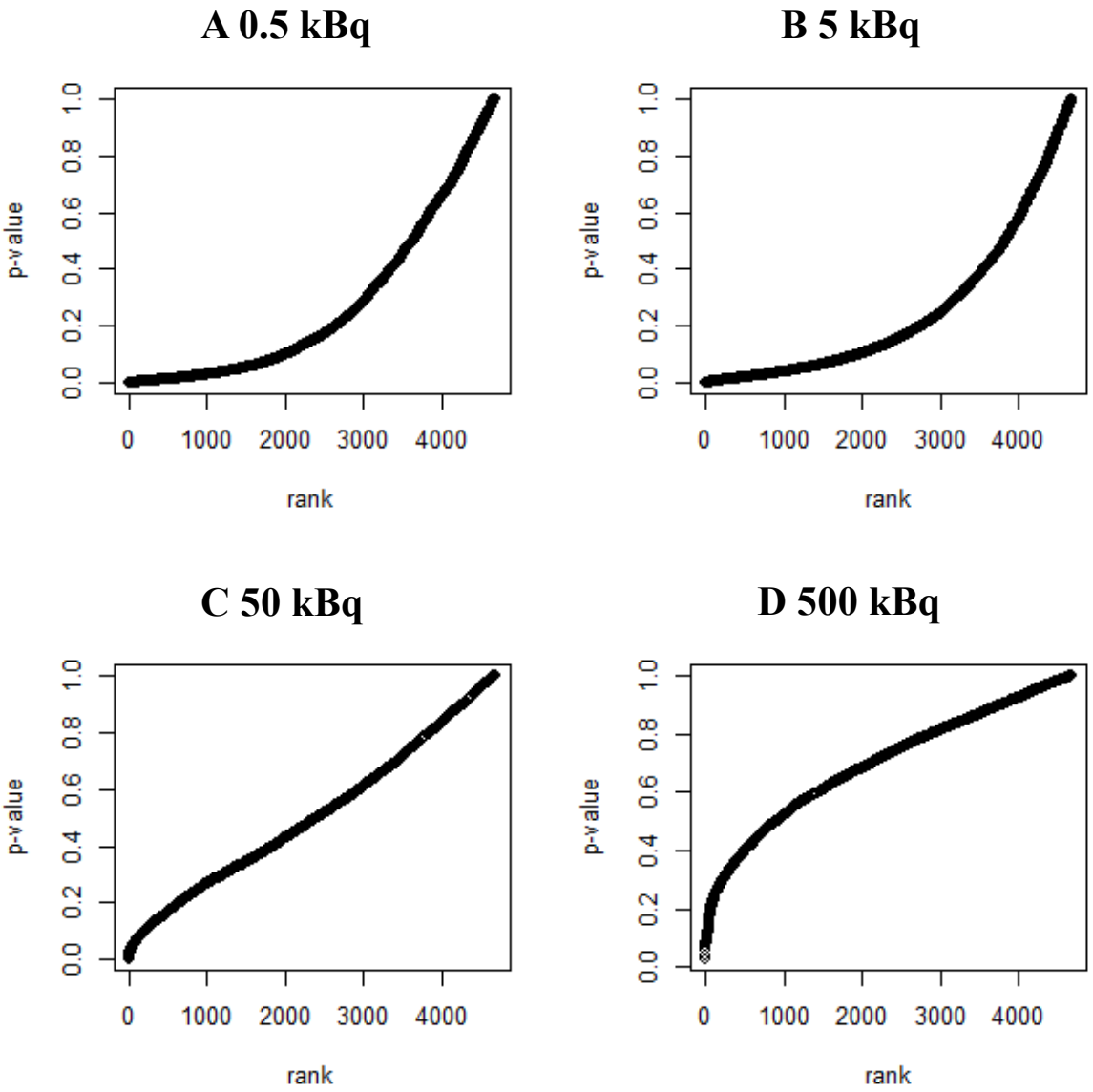

Figure 4: P-values ordered in ascending order plotted against their rank for the young rats at the different investigated activity levels 0.5, 5, 50 and 500 kBq.

Figure 5 is a complex of plots showing the number of proteins for a given FC variance interval for each of the activity groups, 0, 0.5, 5, 50, and 500 kBq, in the two main groups, adult, and young. The plots were truncated to a $\log_2$(FC) variance of 1. All plots contain the same cumulative number of proteins. There were few proteins with a low FC variance in the control and the two highest activity groups in both the young and adult main groups while the activity groups, 0.5 and 5 kBq, had a higher number of low FC variance proteins.

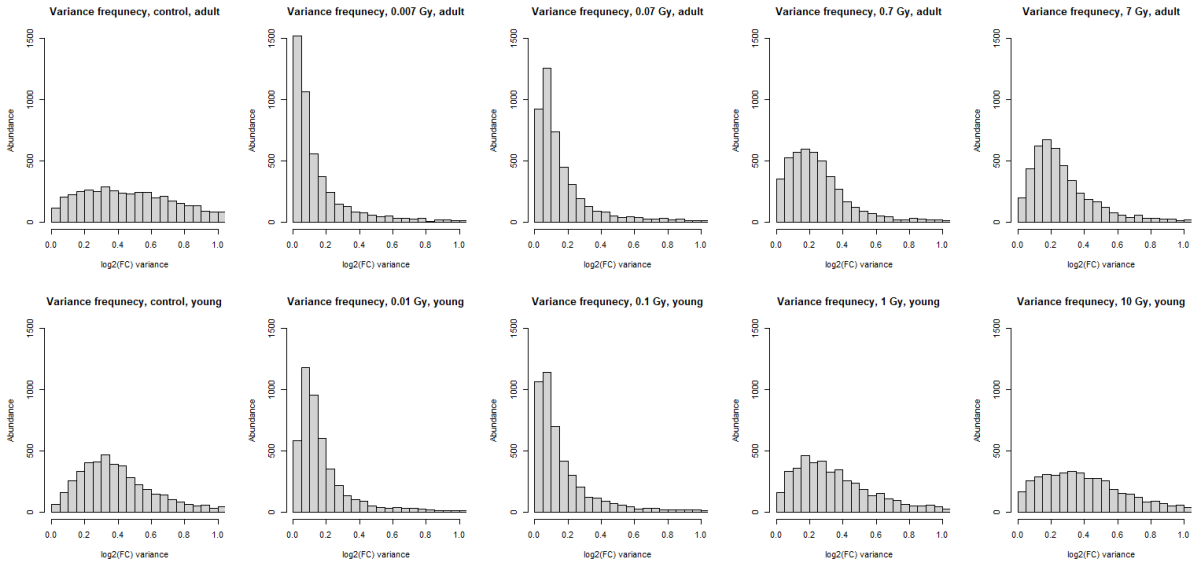

Figure 5: Frequency plots of the groups' log2(FC) variance

Discussion

When looking at Figures 1 and 2 it can be seen that there is a distinct lack of findings except at 0.5 kBq for the young rat population where TRXR3 was found to be the exception. In the cell, TRXR3 is active in the cytosol, mitochondrion, and nucleoplasm. Some biological processes it is involved with are blastocyst formation, cell redox homeostasis, cellular oxidant detoxification, and glutathione metabolic processes. Its functions relates to enabling flavine adenine dinucleotide (FAD) and nicotinamide adenine dinucleotide phosphate (NADP) binding.

The rat protein, TRXR3, is encoded by the *Txnrd3* gene, which is also present in humans. In humans, *TXNRD3* produces the human protein, thioredoxin reductase 3 (TRXR3, denoted here as TRXR3-h for distinguishing between human and rat proteins). TRXR3-h can be found in most of the human body's organs and is expressed the most in the gastrointestinal tract, bone marrow, and lymphoid tissue. TRXR3-h is expressed in the thyroid and more so in the parathyroid. Beyond the places and functions of TRXR3, TRXR3-h can be found in the endoplasmic reticulum and microsomes but is not actively a part of any process in the endoplasmic reticulum or in microsomes. It is also present in cell differentiation and spermatogenesis but is not a part of glutathione metabolic processes.

Research would indicate that insufficient activity of *Txnrd3* leads to cancer inhibition and the presence of TRXR3 prevents inhibition. In a study by Liu, et al., (2022) it was found that knocking out the *Txnrd3* gene led to necrosis from an increased number of reactive oxygen species (ROS) in colon cancer cells, causing increased oxidative stress and therefore inhibiting

tumor proliferation. Liu, et.al, (2022) presented this as reason to why focusing on *Txnrd3* could be useful in treatment of cancer. Other studies indicated that elevated abundance of thioredoxin reductase proteins had a protective effect on the cancer cells by providing control over redox reactions therefore improving cancer cell survival (Patwardhan, et.al, 2022; Sabharwal and Schumacker, 2014). Therefore, the use of TRXR3 inhibitors was proposed as a possible cancer treatment. By blocking the TRXR3 activity the levels of ROS in the tumor cell will increase, causing a higher level of oxidative stress, which could potentially lead to cell death.

The only protein with a statistically significantly altered abundance being connected to redox homeostasis could suggest a possible disruption of redox homeostasis. Since TRXR3 is present in redox activity and redox detoxification, a significant upregulation could suggest an increase in redox activity. In turn, increased redox activity could suggest a higher-than-normal abundance of free radicals or other sources which directly induce oxidation. This could be evidence of increased cell death and cancer occurrence (Branzei and Foiani, 2005; Branzei and Foiani, 2007). However, drawing general conclusions from one protein found at one dose is not advised, but its existence may also point towards something.

No clear dose dependency was observed. The only conclusion Figure 1 and 2 could have offered for dose dependency was decreasing statistical significance for all observed proteins as the dose increased. Other studies (Larsson, et.al, 2020; Larsson, et.al, 2022) would imply that the abundance of some proteins should be significantly increased or decreased at a medium dose as well as low, which is different from this study as no such phenomenon was observed. Hence, these inconsistencies would make it prudent to continue studying dose dependency.

The similarity between the proteins, TRXR3 and TRXR3-h, would suggest that an altered abundance of TRXR3 implies similar implications in humans as in rats. However, the only significant result for TRXR3 was found at 0.5 kBq for the young rats and nowhere else. If the protein cannot consistently be significantly changed in abundance across various doses, it cannot be used as a biomarker for ^{131}I exposure. This also rules out the possibility of using it as a dosimetric biomarker since it needs to not only be consistently changed, but also differentially so. Hence, TRXR3 and therefore TRXR3-h cannot be qualified as a good candidate for biomarking thyroidal ^{131}I exposure or absorbed dose.

It would feel natural to equate overexpression of the gene, *Txnrd3*, with a greater amount of TRXR3, but this would be a mistake. Depending on suppression by translation regulatory proteins, the final amount of TRXR3 becomes uncertain, which could have various implications on cancer proliferation. Indisputable is the fact, however, that an increased level of TRXR3 could still point to oxidative stress caused by increased ROS levels, which is a frequently the case in cancer cells (Patwardhan, et.al, 2022). As such, the presence of high levels of TRXR3 in a tissue could signify cancer, which would make it a theoretical candidate for biomarking cancer.

Analysis

There seemed to have been something which affected the accuracy of the results negatively. When plotting the uncorrected p-values against their rank (Figure 3 and 4), the pattern should ideally have either equal distribution of all ranks among all p-values (suggesting no difference between the control and research groups) or more smaller p-values than higher (suggesting a difference between the control and research groups). In some of the results, the case was a local logarithmic increase. This would imply that the number of low p-values was decreased, i.e., errors had to have been introduced to skew the distribution of p-values.

The error introduction seemed to be dose-dependent. Previous studies (Larsson, et.al, 2020) would suggest statistically significant changes in protein abundance across a wide range of doses, but this does not seem to be the case in this study. When looking at Figure 1 and 2, it seems the statistical accuracy was dose-dependent and that statistically significant changes in protein abundance for low doses became not significant and other proteins' abundances became more insignificant at higher doses. This can also be seen in Figure 3 and 4 as (notably in the young groups) mostly the low-ranking p-values were larger for medium doses than low doses with a prominent logarithmic increase appearing for 500 kBq.

There are similarities between this study and Larsson's (2020). Larsson's study involved a wider pool of analytical methods, containing results from gene expression microarray, proteomic LC-MS/MS, ELISA, and a histological analysis. The proteomic analyses are, however, similar. For instance, the same dose in the same organ for the same age groups were analyzed with the same method. What this and Larsson's study differ in besides the multitude of used methods is the time of euthanasia and final FDR limit. In Larsson's study the time of euthanasia was 9 months, compared to this study's 12 months and the final FDR limit for proteomic analysis was unspecified as samples were pooled and ELISA was used to validate its results.

The similarities in study design and difference analysis method should mean some differences in results. With the difference in time until euthanasia being 3 months, there was time for potential additional changes in protein abundance to have developed within the dose groups of this study beyond that of what Larsson saw. The unspecified FDR limit of the LC-MS/MS analysis does pose potential for differences in results to occur, but it would be unwise to immediately dismiss them seeing as the results were validated through other means. This does still leave the specific FDR to question and means a statistical comparison of numbers becomes irrelevant as Larsson's LC-MS/MS study contains no final numbers. ELISA is, however, a recognized reliable method. Instead of analyzing each rat's thyroid by itself and performing a statistical analysis after, like in this study, Larsson's study pooled each group into one sample and analyzed the pooled sample via LC-MS/MS. This does remove the ability to visualize the variance of each dose group, but again, ELISA is a well-recognized method so the pooling may not play as much of a role as one might think. Therefore, different results would be expected from this study compared to Larsson's, but to what extent is difficult to visualize.

The tissue samples had an amount of additional non-thyroid tissue attached to them, most likely connective tissue. This could have been an introduction of error for some proteins. When analyzed, the extra tissue could have provided protein data common to several types of tissue. Since not all tissue has the same proteome, is equally radiosensitive, or responds in the same way, the additional tissue could have had its own change in protein abundance which interfered with the thyroid's and increased the variance of protein abundance for some proteins in some groups, as seen in Figure 5. Varying responses to various doses coupled with interfering variance of various tissues could have increased the p-value generating the perplexing appearance of Figure 3 and 4. However, this does not completely explain the lack of results as there are some proteins which should not have this problem, like thyroglobulin which is only expressed in the thyroid. Langen et. al (2018) investigated the use of deconvolution to improve statistical accuracy of the method, significance analysis of microarrays. By adding information on tissue composition, the deconvolution algorithm produced a modified output which showed more significant findings compared to no deconvolution. Hence, deconvolution could improve statistical accuracy for proteomic LC-MS/MS.

Not all rats had to have developed any relevant change. As seen in Figure 1 and 2, plenty of proteins were statistically insignificantly abundant relative to the control groups and had an irrelevant change in FC. The implication here was that not all rats had to have developed any changes, since the damage caused by ionizing radiation especially at low doses is stochastic. This would mean that statistical significance on the whole sample population might not be the best way to proceed with an evaluation, as the non-indicative samples could drown the statistics and imply no change. The samples were also not entire thyroids, but bits from thyroids, meaning that if the response was not homogenous within the entire thyroid, the result might not be representative of the entire thyroid proteome. This could have further reduced the probability of finding proteins which might have had altered abundance. Hence, selecting individuals with changes in their proteome and analyzing the entire thyroid might be appropriate. However, as the population of individuals with confirmed change would need to be of a certain size, that size would dictate the total needed number of rats for experimentation which could imply ethical issues.

There are many ways a cell's proteome can be changed and how these changes could contribute to cancer development. One consequence of this is reduced detectability of cancer induction among irradiated groups, as the proteomes of cancer cells could differ. As such, analysis of individuals with confirmed oncogenic changes could be of use in mapping the proteome of these cells, but this does not directly solve the issue of reduced detectability. This data could be used in a comparison of cancer pathways between the young and old populations to find any differences in the type of cancer and carcinogenesis pathways which could be used to improve detectability of other studies through deconvolution (Langen, et.al, 2018). Any other information would have to be supplemented by additional studies, but this way may be a good way to explore proteomic change as studies of the proteome show differing results in which proteins have a significantly changed expression (Larsson, et.al, 2020; Larsson, et.al, 2022).

Proteomic analysis via LC-MS/MS has worked historically for detecting changes in protein abundance after exposure to ^{131}I (Larsson, et.al, 2020; Larsson, et.al, 2022). Probably the greatest strength of proteomic LC-MS/MS is the direct measurement of the proteome, giving an unbiased cross section of concurrent events at the point of measurement. However, variability in the abundance of individual proteins between measurements, also known as batch effect, provide an obstacle to accuracy. To remove these effects is impossible, but they can be suppressed. An example of suppressing this is increased sample sizes, which in turnprovide ethical problems. Resampling techniques, such as bootstrapping (random resampling with replacement) and jackknifing (resampling by systematic omittion), could maybe also be used to improve statistical accuracy to reduce the ethical burden.

Proteomic LC-MS/MS is not without its own problems. During the chromatographic step, the sample is sorted according to, for example size, with some resolution. The resolution ensures that analytes can be differentiated, but the limit of resolution counteracts this and causes analytes similar by the sorting order to be seen as the same. TMT MS/MS techniques quantify the best when only one particle is analysed (Rozanova, et.al, 2021). The practical implications of this would be that proteins are seen as others, which affects annotation and underestimates the peptide quantity (Rozanova, et.al, 2021). Adding more fragmentation to the MS/MS process has shown to counter this problem (Ting, et.al, 2011), but would probably add to the already existing issue of data handling as more information would require processing. In the context of this study and to improve the statistics, an alternative method could be non-destructive confirmation of late effects, followed by an analysis examining different pathways leading to the effects. While not cheap, a possibility is to use nuclear medical imaging techniques to detect the presence of tumour nodules in order to select contenders which would proceed to the next step of analysing their thyroid with other methods, which can be destructive to the thyroid tissue, such as LC-MS/MS or DNA methylation sequencing. Immunoassays, such as ELISA or western blot, could also be used for proteomic studies and would reduce the issues with data handling and quantification, but would require prior knowledge on which proteins should be studied. It should be noted that the entire thyroid would need to be used for analysis to properly catch the affected regions of the thyroid.

The age groups were created to be representative of children and adults giving information on the differences between adult and young rats, but nothing on adolescence. Studying this period and comparing it to the younger and older groups could provide useful thyroid cancer information with regards to thyroid development which could yield knowledge on cancer pathway development by observing when the occurrence of thyroid cancer decreases after internal exposure to ^{131}I. This would add complexity and make it less viable. The use of additional older groups would likely not contribute new information.

Conclusion

One protein with statistically significantly change in relative abundance was found at 0.5 kBq for the young rats. The rat protein, TRXR3, was overexpressed beyond a fold change of 1.5 and is similar in function to the human equivalent, TRXR3, both of which are involved in oxidant detoxification and an over expression would imply the presence of oxidative stress. The implication of oxidative stress could not be confirmed with other significantly regulated proteins. The results pointed towards TRXR3 not being a good candidate biomarker for ^{131}I exposure or absorbed dose due to the lack of appearance across all doses. The high variability within and between groups reduced statistical accuracy, which made it difficult to draw conclusions. Hence more research is needed about variations in response between young and adult individuals after internal irradiation from ^{131}I.

Acknowledgements

The author would like to extend his sincerest gratitude to Eva Forssell Aronsson, Klara Insulander Björk, and Anja Schroff for shouldering the task of supervisors and providing the opportunity for making this work possible, Malin Larsson for allowing the use of her rat tissue samples in this study, Annika Thorsell for helping with troubleshooting the results and providing insight into the LC-MS/MS workflow at Sahlgrenska PCF, Johan Spetz for advising on certain discussion points, Evelin Berger for providing insight into the LC-MS/MS workflow at Sahlgrenska PCF, and Sahlgrenska PCF for providing proteomic LC-MS/MS analysis.

References

Agresti, A. (2012). *Categorical Data Analysis* (3rd ed.). Wiley.

Beckman, D. A., & Feuston, M. (2003). Landmarks in the development of the female reproductive system. *Birth Defects Res B Dev Reprod Toxicol*, *68*(2), 137-143. https://doi.org/10.1002/bdrb.10016

Bell, M. R. (2018). Comparing Postnatal Development of Gonadal Hormones and Associated Social Behaviors in Rats, Mice, and Humans. *Endocrinology*, *159*(7), 2596-2613. https://doi.org/10.1210/en.2018-00220

Berg, G., Jansson, S., Nyström, E., Törring, O., Valdermarsson, S., & Warin, B. (2007). *Tyroideasjukdomar hos vuxna* (E. Nyström, Ed.). Ljungbergs Tryckeri AB Klippan.

Branzei, D., & Foiani, M. (2005). The DNA damage response during DNA replication. *Curr Opin Cell Biol*, *17*(6), 568-575. https://doi.org/10.1016/j.ceb.2005.09.003

Branzei, D., & Foiani, M. (2007). Interplay of replication checkpoints and repair proteins at stalled replication forks. *DNA Repair (Amst)*, *6*(7), 994-1003. https://doi.org/10.1016/j.dnarep.2007.02.018

Bresciani, L., Orlandi, E., & Piazza, C. (2019). Radiation-induced papillary thyroid cancer: is it a distinct clinical entity? *Curr Opin Otolaryngol Head Neck Surg*, *27*(2), 117-122. https://doi.org/10.1097/moo.0000000000000522

Cabanillas, M. E., McFadden, D. G., & Durante, C. (2016). Thyroid cancer. *The Lancet*, *388*(10061), 2783-2795. https://doi.org/https://doi.org/10.1016/S0140-6736(16)30172-6

Cancerregistret. (2022). Cancer i siffror 2018 - Populärvetenskapliga fakta om cancer. In (20/12-2022 ed.). Internet: Socialstyrelsen.

Coleman, M. P., Esteve, J., Damiecki, P., Arslan, A., Renard, H. (1993). Trends in Cancer Incidence and Mortality. In IARC (Ed.), *IARC Scientific Publication* (Vol. 121, pp. 609-640). IARC.

Colevas, A. D., Read, R., Thornhill, J., Adak, S., Tishler, R., Busse, P., Li, Y., & Posner, M. (2001). Hypothyroidism incidence after multimodality treatment for stage III and IV squamous cell carcinomas of the head and neck. *International Journal of Radiation Oncology*Biology*Physics*, *51*(3), 599-604. https://doi.org/https://doi.org/10.1016/S0360-3016(01)01688-1

de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., & Vogel, C. (2009). Global signatures of protein and mRNA expression levels. *Mol Biosyst*, *5*(12), 1512-1526. https://doi.org/10.1039/b908315d

Dunkelmann, S., Wolf, R., Koch, A., Kittner, C., Groth, P., & Schuemichen, C. (2004). Incidence of radiation-induced Graves' disease in patients treated with radioiodine for thyroid autonomy before and after introduction of a high-sensitivity TSH receptor antibody assay. *Eur J Nucl Med Mol Imaging*, *31*(10), 1428-1434. https://doi.org/10.1007/s00259-004-1519-8

EMBL-EBI. (2023). *GO annotations*. EMBL-EBI. Retrieved 19/6 from https://www.ebi.ac.uk/QuickGO/

Fleming, I. D., Black, T. L., Thompson, E. I., Pratt, C., Rao, B., & Hustu, O. (1985). Thyroid dysfunction and neoplasia in children receiving neck irradiation for cancer. *Cancer*, *55*(6), 1190-1194. https://doi.org/10.1002/1097-0142(19850315)55:6<1190::aid-cncr2820550609>3.0.co;2-6

Ghasemi, A., Jeddi, S., & Kashfi, K. (2021). The laboratory rat: Age and body weight matter. *Excli j*, *20*, 1431-1445. https://doi.org/10.17179/excli2021-4072

Hancock, S. L., Cox, R. S., & McDougall, I. R. (1991). Thyroid diseases after treatment of Hodgkin's disease. *N Engl J Med*, *325*(9), 599-605. https://doi.org/10.1056/nejm199108293250902

Hancock, S. L., McDougall, I. R., & Constine, L. S. (1995). Thyroid abnormalities after therapeutic external radiation. *Int J Radiat Oncol Biol Phys*, *31*(5), 1165-1170. https://doi.org/10.1016/0360-3016(95)00019-u

Howlader, N., Noone, A., Krapcho, M., Miller, D., Bishop, K., Altekruse, S., Kosary, C., Yu, M., Ruhl, J., & Tatalovich, Z. (2016). SEER cancer statistics review, 1975-2013, national cancer institute. bethesda, MD. In.

Illés, Á., Bíró, E., Miltényi, Z., Keresztes, K., Váróczy, L., András, C., Sipka, S., & Bakó, G. (2003). Hypothyroidism and Thyroiditis after Therapy for Hodgkin's Disease. *Acta Haematologica*, *109*(1), 11-17. https://doi.org/10.1159/000067269

Imaizumi, M., Ohishi, W., Nakashima, E., Sera, N., Neriishi, K., Yamada, M., Tatsukawa, Y., Takahashi, I., Fujiwara, S., Sugino, K., Ando, T., Usa, T., Kawakami, A., Akahoshi, M., & Hida, A. (2017). Thyroid Dysfunction and Autoimmune Thyroid Diseases Among Atomic Bomb Survivors Exposed in Childhood. *J Clin Endocrinol Metab*, *102*(7), 2516-2524. https://doi.org/10.1210/jc.2017-00102

Kailasiam, S. (2021). LC-MS – What Is LC-MS, LC-MS Analysis and LC-MS/MS. *Technology Networks*. https://www.technologynetworks.com/analysis/articles/lc-ms-what-is-lc-ms-lc-ms-analysis-and-lc-msms-348238

Langen, B., Rudqvist, N., Spetz, J., Helou, K., & Forssell-Aronsson, E. (2018). Deconvolution of expression microarray data reveals 131I-induced responses otherwise undetected in thyroid tissue. *PLoS One*, *13*(7), e0197911. https://doi.org/10.1371/journal.pone.0197911

Larsson, M., Rudqvist, N., Spetz, J., Shubbar, E., Parris, T. Z., Langen, B., Helou, K., & Forssell-Aronsson, E. (2020). Long-term transcriptomic and proteomic effects in Sprague Dawley rat thyroid and plasma after internal low dose 131I exposure. *PLoS One*, *15*(12), e0244098. https://doi.org/10.1371/journal.pone.0244098

Larsson, M., Rudqvist, N. P., Spetz, J., Parris, T. Z., Langen, B., Helou, K., & Forssell-Aronsson, E. (2022). Age-related long-term response in rat thyroid tissue and plasma after internal low dose exposure to (131)I. *Sci Rep*, *12*(1), 2107. https://doi.org/10.1038/s41598-022-06071-4

Liu, Q., Du, P., Zhu, Y., Zhang, X., Cai, J., & Zhang, Z. (2022). Thioredoxin reductase 3 suppression promotes colitis and carcinogenesis via activating pyroptosis and necrosis. *Cell Mol Life Sci*, *79*(2), 106. https://doi.org/10.1007/s00018-022-04155-y

Maier, T., Güell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Lett*, *583*(24), 3966-3973. https://doi.org/10.1016/j.febslet.2009.10.036

Marty, M. S., Chapin, R. E., Parks, L. G., & Thorsrud, B. A. (2003). Development and maturation of the male reproductive system. *Birth Defects Res B Dev Reprod Toxicol*, *68*(2), 125-136. https://doi.org/10.1002/bdrb.10015

Mattsson, S., Johansson, L., Leide Svegborn, S., Liniecki, J., Noßke, D., Riklund, K. Å., Stabin, M., Taylor, D., Bolch, W., Carlsson, S., Eckerman, K., Giussani, A., Söderberg, L., & Valind, S. (2015). ICRP Publication 128: Radiation Dose to Patients from Radiopharmaceuticals: a Compendium of Current Information Related to Frequently Used Substances. *Annals of the ICRP*, *44*(2_suppl), 7-321. https://doi.org/10.1177/0146645314558019

Mercado, G., Adelstein, D. J., Saxton, J. P., Secic, M., Larto, M. A., & Lavertu, P. (2001). Hypothyroidism [https://doi.org/10.1002/1097-0142(20011201)92:11<2892::AID-CNCR10134>3.0.CO;2-T]. *Cancer*, *92*(11), 2892-2897. https://doi.org/https://doi.org/10.1002/1097-0142(20011201)92:11<2892::AID-CNCR10134>3.0.CO;2-T

Metso, S., Jaatinen, P., Huhtala, H., Luukkaala, T., Oksala, H., & Salmi, J. (2004). Long-term follow-up study of radioiodine treatment of hyperthyroidism. *Clin Endocrinol (Oxf)*, *61*(5), 641-648. https://doi.org/10.1111/j.1365-2265.2004.02152.x

Milton J.S., A. J. C. (1995). *Introduction to Probability and Statistics, Principles and Applications for Engineering and the Computing Sciences* (B. J. W. Shira J., Ed. 3rd ed.). McGraw-Hill.

Nagayama, Y. (2018). Radiation-related thyroid autoimmunity and dysfunction. *J Radiat Res*, *59*(suppl_2), ii98-ii107. https://doi.org/10.1093/jrr/rrx054

NCI. (2023, Unknown). *Thyroid Cancer - Health Professional Version*. National Cancer Institute. Retrieved 24/1 from https://www.cancer.gov/types/thyroid/hp

Nygaard, B., Knudsen, J. H., Hegedüs, L., Scient, A. V., & Hansen, J. E. (1997). Thyrotropin receptor antibodies and Graves' disease, a side-effect of 131I treatment in patients with nontoxic goiter. *J Clin Endocrinol Metab*, *82*(9), 2926-2930. https://doi.org/10.1210/jcem.82.9.4227

Patwardhan, R. S., Sharma, D., & Sandur, S. K. (2022). Thioredoxin reductase: An emerging pharmacologic target for radiosensitization of cancer. *Transl Oncol*, *17*, 101341. https://doi.org/10.1016/j.tranon.2022.101341

Picut, C. A., Ziejewski, M. K., & Stanislaus, D. (2018). Comparative Aspects of Pre- and Postnatal Development of the Male Reproductive System. *Birth Defects Res*, *110*(3), 190-227. https://doi.org/10.1002/bdr2.1133

Quinn, R. (2005). Comparing rat's to human's age: how old is my rat in people years? *Nutrition*, *21*(6), 775-777. https://doi.org/10.1016/j.nut.2005.04.002

Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H., & Marcus, K. (2021). Quantitative Mass Spectrometry-Based Proteomics: An Overview. *Methods Mol Biol*, *2228*, 85-116. https://doi.org/10.1007/978-1-0716-1024-4_8

Sabharwal, S. S., & Schumacker, P. T. (2014). Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? *Nature Reviews Cancer*, *14*(11), 709-721. https://doi.org/10.1038/nrc3803

Semple, B. D., Blomgren, K., Gimlin, K., Ferriero, D. M., & Noble-Haeusslein, L. J. (2013). Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol*, *106-107*, 1-16. https://doi.org/10.1016/j.pneurobio.2013.04.001

Spetz, J., Rudqvist, N., & Forssell-Aronsson, E. (2013). Biodistribution and dosimetry of free 211At, 125I- and 131I- in rats. *Cancer Biother Radiopharm*, *28*(9), 657-664. https://doi.org/10.1089/cbr.2013.1483

Stanley, D. P., & Shetty, A. K. (2004). Aging in the rat hippocampus is associated with widespread reductions in the number of glutamate decarboxylase-67 positive interneurons but not interneuron degeneration. *J Neurochem*, *89*(1), 204-216. https://doi.org/10.1111/j.1471-4159.2004.02318.x

Thomas, G., & Yamashita, S. (2017). Thirty Years After Chernobyl and 5 After Fukushima—What Have We Learnt and What Do We Still Need to Know? In S. Yamashita & G. Thomas (Eds.), *Thyroid Cancer and Nuclear Accidents: Long-Term Aftereffects of Chernobyl and Fukushima* (pp. xv-xxiv). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-812768-1.00026-5

Ting, L., Rad, R., Gygi, S. P., & Haas, W. (2011). MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods*, *8*(11), 937-940. https://doi.org/10.1038/nmeth.1714

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., . . . Pontén, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science*, *347*(6220), 1260419. https://doi.org/10.1126/science.1260419

UNSCEAR. (2017). *Evaluation of data on thyroid cancer in regions affected by the Chernobyl accident* (White Paper, Issue. UNSCEAR. https://www.unscear.org/docs/publications/2017/Chernobyl_WP_2017.pdf

Vidal, J. D. (2017). The Impact of Age on the Female Reproductive System. *Toxicol Pathol*, *45*(1), 206-215. https://doi.org/10.1177/0192623316673754

Vogelius, I. R., Bentzen, S. M., Maraldo, M. V., Petersen, P. M., & Specht, L. (2011). Risk factors for radiation-induced hypothyroidism: a literature-based meta-analysis. *Cancer*, *117*(23), 5250-5260. https://doi.org/10.1002/cncr.26186

Zhang, L., & Elias, J. E. (2017). Relative Protein Quantification Using Tandem Mass Tag Mass Spectrometry. In L. Comai, J. E. Katz, & P. Mallick (Eds.), *Proteomics: Methods and Protocols* (pp. 185-198). Springer New York. https://doi.org/10.1007/978-1-4939-6747-6_14

APPENDIX

R Code

```
install.packages("tidyverse")
install.packages("readxl")
install.packages("writexl")

if (!requireNamespace('BiocManager', quietly = TRUE))
   install.packages('BiocManager')
BiocManager::install('EnhancedVolcano')

library(tidyverse)
library(readxl)
library(writexl)
library(knitr)
library(EnhancedVolcano)

setwd("") # Set work directory

rm(list = ls()) # Delets the current workspace for easier debugging

mydata <- read_xlsx("", sheet = "") # Select excel file and work sheet
mydata <- mydata %>% drop_na

exprdata <- mydata %>% select(c(X1:N4))
boxplot(exprdata, ylab="Relativee fold change", xlab="Group")

# normalcy check
hist(unlist(exprdata),breaks=100)

exprdata_log <- log2(exprdata)
hist(unlist(exprdata_log),breaks=100)

exprdata_log_t <- t(exprdata_log)

# principal component analysis
pc <- prcomp(exprdata_log_t, center=TRUE, scale=TRUE)
plot(pc)

pc_comp <- pc$x

col1="blue"
col2="green"
col3="yellow"
col4="orange"
col5="red"

# principal component plots, same dose has same colour
plot(pc_comp[1:nrow(pc_comp),1], pc_comp[1:nrow(pc_comp),2], col=rep(c("black", each=1)),
xlab="PC1", ylab="PC2", pch=16)

X=1:5
text(pc_comp[X,1], pc_comp[X,2], colnames(exprdata_log)[X], col=col1, pos=2)

I=6:10
text(pc_comp[I,1], pc_comp[I,2], colnames(exprdata_log)[I], col=col2, pos=2)

H=11:16
text(pc_comp[H,1], pc_comp[H,2], colnames(exprdata_log)[H], col=col3, pos=2)
```

```
O=17:21
text(pc_comp[O,1], pc_comp[O,2], colnames(exprdata_log)[O], col=col4, pos=2)

R=22:27
text(pc_comp[R,1], pc_comp[R,2], colnames(exprdata_log)[R], col=col5, pos=2)

U=28:33
text(pc_comp[U,1], pc_comp[U,2], colnames(exprdata_log)[U], col=col1, pos=4)

E=34:38
text(pc_comp[E,1], pc_comp[E,2], colnames(exprdata_log)[E], col=col2, pos=4)

D=39:42
text(pc_comp[D,1], pc_comp[D,2], colnames(exprdata_log)[D], col=col3, pos=4)

C=43:47
text(pc_comp[C,1], pc_comp[C,2], colnames(exprdata_log)[C], col=col4, pos=4)

N=48:51
text(pc_comp[N,1], pc_comp[N,2], colnames(exprdata_log)[N], col=col5, pos=4)

exprdata_out <- cbind(Accession = mydata$Accession,
                exprdata,
                mean_Control_y = apply(exprdata[,U], 1, mean),
                mean_Control_o = apply(exprdata[,X], 1, mean),
                                                mean_001Gy_y = apply(exprdata[,E], 1, mean),
                                                mean_001Gy_o = apply(exprdata[,I], 1, mean),
                                                mean_01Gy_y = apply(exprdata[,D], 1, mean),
                                                mean_01Gy_o = apply(exprdata[,H], 1, mean),
                                                mean_1Gy_y = apply(exprdata[,C], 1, mean),
                                                mean_1Gy_o = apply(exprdata[,O], 1, mean),
                                                mean_10Gy_y = apply(exprdata[,N], 1, mean),
                                                mean_10Gy_o = apply(exprdata[,R], 1, mean))
head(exprdata_out)

exprdata_out <- exprdata_out %>% mutate(log2FC_001Gy_y = log2(mean_001Gy_y/mean_Control_y))
exprdata_out <- exprdata_out %>% mutate(log2FC_001Gy_o = log2(mean_001Gy_o/mean_Control_o))
exprdata_out <- exprdata_out %>% mutate(log2FC_01Gy_y = log2(mean_01Gy_y/mean_Control_y))
exprdata_out <- exprdata_out %>% mutate(log2FC_01Gy_o = log2(mean_01Gy_o/mean_Control_o))
exprdata_out <- exprdata_out %>% mutate(log2FC_1Gy_y = log2(mean_1Gy_y/mean_Control_y))
exprdata_out <- exprdata_out %>% mutate(log2FC_1Gy_o = log2(mean_1Gy_o/mean_Control_o))
exprdata_out <- exprdata_out %>% mutate(log2FC_10Gy_y = log2(mean_10Gy_y/mean_Control_y))
exprdata_out <- exprdata_out %>% mutate(log2FC_10Gy_o = log2(mean_10Gy_o/mean_Control_o))
head(exprdata_out)

# t-tests for each group compared to the controll
exprdata_out <- cbind(exprdata_out,
                pvalue_001_y = apply(exprdata_log, 1, function(x) {t.test(x[E], x[U])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_001_o = apply(exprdata_log, 1, function(x) {t.test(x[I], x[X])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_01_y = apply(exprdata_log, 1, function(x) {t.test(x[D], x[U])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_01_o = apply(exprdata_log, 1, function(x) {t.test(x[H], x[X])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_1_y = apply(exprdata_log, 1, function(x) {t.test(x[C], x[U])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_1_o = apply(exprdata_log, 1, function(x) {t.test(x[O], x[X])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_10_y = apply(exprdata_log, 1, function(x) {t.test(x[N], x[U])$p.value}))
exprdata_out <- cbind(exprdata_out,
                pvalue_10_o = apply(exprdata_log, 1, function(x) {t.test(x[R], x[X])$p.value}))

exprdata_out <- cbind(exprdata_out, fdr_001_y = p.adjust(exprdata_out$pvalue_001_y, method="BH"))
```

```
exprdata_out <- cbind(exprdata_out, fdr_001_o = p.adjust(exprdata_out$pvalue_001_o, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_01_y = p.adjust(exprdata_out$pvalue_01_y, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_01_o = p.adjust(exprdata_out$pvalue_01_o, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_1_y = p.adjust(exprdata_out$pvalue_1_y, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_1_o = p.adjust(exprdata_out$pvalue_1_o, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_10_y = p.adjust(exprdata_out$pvalue_10_y, method="BH"))
exprdata_out <- cbind(exprdata_out, fdr_10_o = p.adjust(exprdata_out$pvalue_10_o, method="BH"))

alfa=0.05
FC=log2(1.5)
labelsize = 0 # makes dots visible while text becomes minimal
axislabelsize = 20
titlesize = 30
x=c(-3,3)
y=c(0,1.7)

sortorder001y = sort(exprdata_out$pvalue_001_y)
sortorder001o = sort(exprdata_out$pvalue_001_o)
sortorder01y = sort(exprdata_out$pvalue_01_y)
sortorder01o = sort(exprdata_out$pvalue_01_o)
sortorder1y = sort(exprdata_out$pvalue_1_y)
sortorder1o = sort(exprdata_out$pvalue_1_o)
sortorder10y = sort(exprdata_out$pvalue_10_y)
sortorder10o = sort(exprdata_out$pvalue_10_o)

# plots for analysis of accuracy
mylayout = matrix(1:4,2,2,byrow = TRUE)
layout(mylayout)
plot(1:length(sortorder001y),sortorder001y, xlab = "rank", ylab = "p-value", main = "0.01 Gy, young")
plot(1:length(sortorder01y),sortorder01y, xlab = "rank", ylab = "p-value", main = "0.1 Gy, young")
plot(1:length(sortorder1y),sortorder1y, xlab = "rank", ylab = "p-value", main = "1 Gy, young")
plot(1:length(sortorder10y),sortorder10y, xlab = "rank", ylab = "p-value", main = "10 Gy, young")
win.graph()
mylayout = matrix(1:4,2,2,byrow = TRUE)
layout(mylayout)
plot(1:length(sortorder001o),sortorder001o, xlab = "rank", ylab = "p-value", main = "0.007 Gy, adult")
plot(1:length(sortorder01o),sortorder01o, xlab = "rank", ylab = "p-value", main = "0.07 Gy, adult")
plot(1:length(sortorder1o),sortorder1o, xlab = "rank", ylab = "p-value", main = "0.7 Gy, adult")
plot(1:length(sortorder10o),sortorder10o, xlab = "rank", ylab = "p-value", main = "7 Gy, adult")

# volcano plots of BH corrected p-values over FC
win.graph()
EnhancedVolcano(exprdata_out,                    # data to plot
        lab = exprdata_out$Accession,       # labels (protein name)
        x = "log2FC_001Gy_y",                    # value to plot in the X-axis
        y = "fdr_001_y",                  # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "0.01 Gy, young",       # Title of the plot
        subtitle = NULL,                 # Remove subtitle to create more space for the plot
        caption = NULL,                   # Remove caption to create more space for the plot
                                titleLabSize = titlesize,
                                # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                 # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                 # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                    # data to plot
        lab = exprdata_out$Accession,       # labels (protein name)
        x = "log2FC_01Gy_y",                    # value to plot in the X-axis
        y = "fdr_01_y",                  # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
```

```
        FCcutoff = FC,   # cut off for the fold change
        title = "0.1 Gy, young",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_1Gy_y",                    # value to plot in the X-axis
        y = "fdr_1_y",                   # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "1 Gy, young",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_10Gy_y",                    # value to plot in the X-axis
        y = "fdr_10_y",                   # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "10 Gy, young",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_001Gy_o",                    # value to plot in the X-axis
        y = "fdr_001_o",                   # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "0.007 Gy, adult",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
```

```
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_01Gy_o",                  # value to plot in the X-axis
        y = "fdr_01_o",                    # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "0.07 Gy, adult",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_1Gy_o",                   # value to plot in the X-axis
        y = "fdr_1_o",                     # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "0.7 Gy, adult",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7
win.graph()
EnhancedVolcano(exprdata_out,                     # data to plot
        lab = exprdata_out$Accession,        # labels (protein name)
        x = "log2FC_10Gy_o",                  # value to plot in the X-axis
        y = "fdr_10_o",                    # value to plot in the Y-axis
        pCutoff = alfa,    # cut off for the pvalue
        FCcutoff = FC,   # cut off for the fold change
        title = "7 Gy, adult",       # Title of the plot
        subtitle = NULL,                  # Remove subtitle to create more space for the plot
        caption = NULL,                    # Remove caption to create more space for the plot
                                  titleLabSize = titlesize,
                                  # (if you remove this line you will get the number of proteins plotted)
        legendPosition = "top",             # Position the legend on top of the plot
        axisLabSize = axislabelsize,                # Set font size for axis labels
        labSize = labelsize,                 # Set font size for protein labels
        xlim = x,                  # Set x-axis limits to -3 and 3 so the plot is symmetric
        ylim = y)                  # Set y-axis limits to 0 and 7


log2varX = diag(as.matrix(var(t(exprdata_log[,X]))))
log2varI = diag(as.matrix(var(t(exprdata_log[,I]))))
log2varH = diag(as.matrix(var(t(exprdata_log[,H]))))
log2varO = diag(as.matrix(var(t(exprdata_log[,O]))))
log2varR = diag(as.matrix(var(t(exprdata_log[,R]))))
log2varU = diag(as.matrix(var(t(exprdata_log[,U]))))
log2varE = diag(as.matrix(var(t(exprdata_log[,E]))))
log2varD = diag(as.matrix(var(t(exprdata_log[,D]))))
log2varC = diag(as.matrix(var(t(exprdata_log[,C]))))
log2varN = diag(as.matrix(var(t(exprdata_log[,N]))))

vardata <- cbind(Accession = mydata$Accession,
                 log2varX,log2varI,log2varH,log2varO,log2varR,
```

```
                log2varU,log2varE,log2varD,log2varC,log2varN)
vardata = as.data.frame(vardata)

x = c(0,1)
y = c(0,1500)
mybreaks = 400
mytext = "Variance frequnecy,"
xname = "log2(FC) variance"
yname = "Abundance"
mylayout = matrix(1:10,2,5,byrow = TRUE)
layout(mylayout)

# Plotting of variance frequency historgrams
win.graph()
hist(as.numeric(vardata$log2varX),
                breaks = mybreaks,
                main = paste(mytext,"control, adult"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varI),
                breaks = mybreaks,
                main = paste(mytext,"0.007 Gy, adult"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varH),
                breaks = mybreaks,
                main = paste(mytext,"0.07 Gy, adult"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varO),
                breaks = mybreaks,
                main = paste(mytext,"0.7 Gy, adult"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varR),
                breaks = mybreaks,
                main = paste(mytext,"7 Gy, adult"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varU),
                breaks = mybreaks,
                main = paste(mytext,"control, young"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varE),
                breaks = mybreaks,
                main = paste(mytext,"0.01 Gy, young"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varD),
```

```
                breaks = mybreaks,
                main = paste(mytext,"0.1 Gy, young"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varC),
                breaks = mybreaks,
                main = paste(mytext,"1 Gy, young"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
hist(as.numeric(vardata$log2varN),
                breaks = mybreaks*2,
                main = paste(mytext,"10 Gy, young"),
                xlab = xname,
                ylab = yname,
                xlim = x,
                ylim = y)
```