The University of Birmingham

School of Biosciences

Project 1

MOLECULAR MECHANISMS OF NEUROTOXICITY AND GENOTOXICITY BY BROMINATED FLAME RETARDANTS IN SH-SY5Y HUMAN NEUROBLASTOMA CELLS

Project 2

WHOLE-GENOME SEQUENCING, FINISHING AND ANALYSIS OF THREE BACTERIAL HUMAN PATHOGENS

By

Jelena Sostare

A combined research thesis submitted to the University of Birmingham as a part of the requirement for the degree of MRes in Molecular and Cellular Biology

Birmingham, 2013

Summary

This combined research thesis, submitted to the University of Birmingham, consists of two projects. The first project addressed the neurotoxic effects by three of the most widely used flame retardants – hexabromocylododecane, tetrabromobisphenol-A and decabromodiphenyl ether in SH-SY5Y human neuroblastoma cells. The results demonstrated high toxicity potential in all three compounds even at low concentrations. Moreover, all compounds caused DNA single-strand breaks at non-cytotoxic concentrations. HBCD proved to be more potent than other two compounds tested. Therefore, it can be concluded that all three compounds are potentially neurotoxic and what more, genotoxic in human cells *in-vitro.*

The second project attempted to finish the genome of multidrug-resistant *Elizabethkingia meningoseptica* 501 and start finishing the genome of a new *Pseudomonas aeruginosa* ST395 strain, employing whole-genome Nextera XT and Mate Pair sequencing (Illumina). Optimization of Nextera XT for organisms with different GC content was also carried out in the study using 3 species – *Escherichia coli* (Medium GC), *Pseudomonas aeruginosa* (high GC) and *Elizabethkingia meningoseptica* (low GC). Finally, using PCR and Sanger sequencing, the genome of *E. meningoseptica* was almost completed, obtaining 4 gapped fragments of the genome, what makes HTS (high-throughput sequencing) a very powerful scientific tool.

The University of Birmingham

School of Biosciences

MOLECULAR MECHANISMS OF NEUROTOXICITY AND GENOTOXICITY BY BROMINATED FLAME RETARDANTS IN SH-SY5Y HUMAN NEUROBLASTOMA CELLS

A research project report submitted by

Jelena Sostare

as a part of the requirement for the

degree of MRes in Molecular and Cellular Biology

Project supervisors: Dr Nik J. Hodges, Dr Frank Michelangeli

Birmingham, 2013

Abstract

Brominated flame retardants (BFR) are substances, added to various consumer products to prevent their inflammation. They are widespread pollutants in the environment and ubiquitous contaminants in humans. So far only few studies assessed the neurotoxicity of brominated flame retardants, therefore in this report neurotoxic effects by three of the most widely used flame retardants – hexabromocylododecane (HBCD) tetrabromobisphenol-A (TBBP-A) and decabromodiphenyl ether (BDE-209) in SH-SY5Y human neuroblastoma cells were investigated. The results demonstrated high toxicity potential in all three compounds even at low concentrations (with LC_{50} values between 2.5 and 5 µM, 15 and 20 µM, 30 and 60 µM for HBCD, TBBP-A and BDE-209, respectively). Concentration-dependent increases in reactive oxygen species (ROS) were observed in cells treated with TBBP-A, but not in cells treated with HBCD. There was also a depletion of neuronal antioxidant – glutathione (GSH) following treatment of cells with HBCD, but no response with TBBP-A. In contrast, BDE-209 caused only minor changes in ROS and GSH levels. Following treatment with HBCD, TBBP-A and BDE-209 there was increased annexin V staining, indicative of apoptosis. HBCD treated cells also showed increased PI labeling (necrosis indicator) and autophagic puncta in GFP transfected cells. Furthermore, DNA strand breaks were also measured. It has been known that BDE-209 causes DNA strand breaks in neuronal cells, but there have been no studies investigating the ability of HBCD and TBBP-A to induce DNA damage. In the current study, all compounds caused DNA single-strand breaks at non-cytotoxic concentrations. HBCD proved to be more potent than other two compounds tested. In summary, it can be concluded that all three compounds are potentially neurotoxic and what more, genotoxic in human cells *in-vitro.*

Acknowledgements

I would like to acknowledge Dr Nik J. Hodges and Dr Frank Michelangeli for excellent supervising and support. I also would like to thank Louise Stone, Lorna Thorne, Sunil Claire, Shrikant Jondhale, Mohammed Shuwaikan and Zahra Khan for helping in the lab.

Contents

Chart list

List of figures

Abbreviations

ABS	Acrylonitrile-butadiene-styrene
ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
BDE	Brominated diphenyl ether
BDE-47	Tetrabromodiphenyl ether
BDE-209, DBPE or decaBDE	Decabromodiphenyl ether
BFR	Brominated flame retardants
BSA	Bovine serum albumin
Ca2+	Calcium
DCF	2', 7' –dichlorofluorescin
DCFDA	2',7' –dichlorofluorescein diacetate
dH_2O	Distilled water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimhethyl sulfoxide
DNA	Deoxyribonucleic acid
EPS	Expanded polysterene
EU	European Union
FADD	Fas-Associated Death Domain protein
FITC	Fluorescein isothiocyanate
FLIP	FLICE-like inhibitory protein
FR	Flame retardants
GFP	Green Fluorescent Protein
GSH	Glutathione
GSSG	Glutathione disulfide
H_2O_2	Hydrogen peroxide
HBCD	Hexabromocylododecane
KFERQ	Amino acid sequence Lys-Phe-Glu-Arg-Gln
LB	Lysogeny broth
LC3	Rat microtubule-associated protein 1 light chain 3
LC50	Lethal concentration 50%
LMPA	Low melting point agarose
MTT	3-[4,5-Dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide

NMPA	Normal melting point agarose
NOEL	No Observable Effect Level
OPT	O-Phthalaldehyde
PBDE	Polybrominated diphenyl ethers
PCD	Programmed cell death
pEGFP-C1	Enhanced green fluorescent protein plasmid
PI	Propidium iodide
POPs	Persistent organic pollutants
RIP1	Receptor-interacting protein 1
RNS	Reactive nitrogen species
ROS	Reactive Oxygen Species
SH-SY5Y	Human neuroblastoma cell line - differentiated
SOD	Superoxide dismutase
SYBR	Synergy Brands, Inc
TBBP-A	Tetrabromobisphenol A
TCA	Trichloroacetic Acid
TTR	Transthyretin
UV	Ultraviolet

1. Introduction

1.1. Brominated flame retardants

In the modern world production of electronical appliances, materials for building construction and domestic use is rising every year and along with those the demand for Brominated flame retardants (BFR) is increasing (Alaee et al., 2003). While in 1990 only 145 000 metric tons of BFRs were produced worldwide (Pettigrew, 1994), by 2000 the production increased almost 2 fold making 310 000 tones (BSEF, 2000), and in 2008 the number reached 410 000 tons (Fink et al., 2008; **Chart 1.1**).

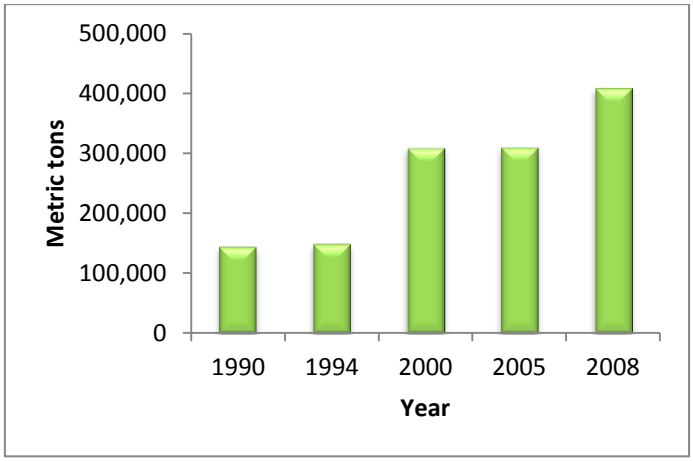


Chart 1.1. BFR production form 1990 – 2008, metric tons.

Brominated flame retardants are substances that lower the risk of ignition in electronics, furnishing and construction materials. They are chemicals with long degradation period widely spread in the environment, and as a result BFRs are referred as persistent organic pollutants (POPs) (Al-Mousa & Michelangeli, 2012). Brominated flame retardants are found not only close to the production sources, but also in regions distant from the source as they may spread via air or water (Fonnum & Mariussen, 2009). These pollutants can be bioaccamulated by animals and humans, causing neurological disorders and influencing

learning, memory and motor functions (Al-Mousa & Michelangeli, 2012). BFR influence on nervous system is of particular concern as neuronal cells are very sensitive to toxicity especially in early development period (Weiss 2000).

At the present time there are more than 175 compounds in the flame retardant (FR) group. FRs are represented by 4 main families: inorganic (50% of all FRs per year) such as aluminium trihydroxide and red phosphorus; halogenated organic (25% of all FRs per year) chlorine and bromine (BFR) based; organophosphorus (20% of all FRs per year), mainly phosphate esters and the rarest - nitrogen-based (>5% of all FRs per year) (EHC-192, 1997).

According to the way BFRs create polymers they are classified as brominated monomers, reactive or additive. From brominated monomers, for instance, brominated styrene or butadiene complex compounds can be created. Reactive BFRs, for example, tetrabromobisphenol A (TBBP-A) are added to polymer products and are embedded into them. Whereas additive flame retardants such as polybrominated diphenyl ethers (PBDEs) and hexabromocylododecane (HBCD) are mixed with the manufactured product without chemical bonding and they have higher probability to leach into the environment (Hutzinger & Thoma, 1987).

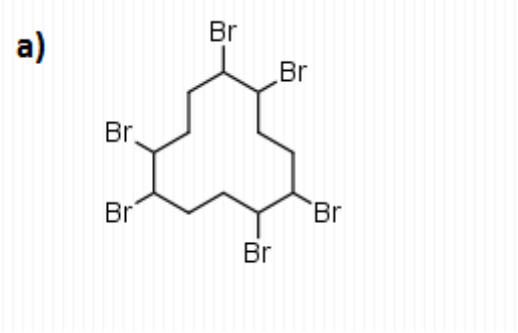
1.2. HBCD

Hexabromocyclododecanes (HBCDs) are brominated cycloaliphatic additive hydrocarbons used to diminish flammability in construction materials, upholstery textiles and electronical devices. At the present moment HBCDs are referred as persistent, toxic and ubiquitous contaminants since they are common in the environment and bioaccumulative in humans (Covaci et al., 2006).

1.2.1. Chemical and physical characteristics

The compound 1,2,5,6,9,10-hexabromocyclododecane (HBCD, $C_{12}H_{18}Br_6$) is a white crystalline powder, with 74.7% bromine (**Figure 1.1. a**). It is thermally rearranged at temperatures above 160 °C, what leads to formation of 3 specific diastereoisomers – α, β and γ. The most common γ isomer makes up to 70% of the final HBCD mixture, while α and β are represented only by 10 and 6% (the rest are additives; Marvin et al., 2011). The difference in isomer structure may cause various polarities, dipole moments and water solubility. Solubility for γ, α and β isomers was measured by Hunziker et al. accounting 48.8, 14.7 and 2.1 íg/L, respectively (Hunziker et al 2004). These dissimilarities may lead to different biological uptake, metabolism and could explain diverse environmental behaviour (de Wit, 2002; Law et al., 2003).

All previous studies proved that diastereoisomer γ dominates only in technical HBCD and along with trophic level shift (food chain development from primary producers to predators) accumulation of isomer α is increasing and becomes greater than γ and β (Law et al., 2006; Law et al., 2007; Covaci et al., 2006; Tomy et al., 2008). This shift through the food chain is possibly a consequence of varied water solubility, bioavailability, uptake and depuration kinetics (including biotransformation) and better *in-vivo* bioisomerization of α-HBCD compared with 2 other isomers (Covaci et al., 2006; Wu et al., 2010).



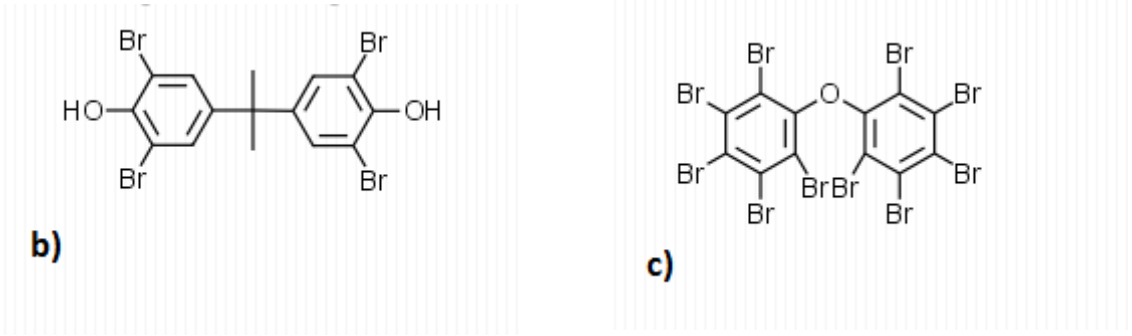


Figure 1.1. General structure of a) HBCD, b) TBBP-A, c) BDE-209 (Created using ChemDoodle software).

1.2.2. Occurrence and exposure

The production of HBCD started in 1960s and today it is one of the most commonly used brominated flame retardant. In 2001 HBCD was taking the 3rd place in annual production of all BFRs (**Chart 1.2**) in the world - 16 700 t, after only tetrabromobisphenol-A (>130 000 t) and decabromodiphenyl ether (56 100 t). EU consumption of the compound in 2005 was 11 000 t (Janak et al., 2005), what is almost half of produced HBCD. Generally HBCD is used in construction industry where is included approximately at <3% by weight into polystyrene foam materials. It is also employed in production of upholstered furniture, transport interior textiles and cushions, electric and electronic devices (Marvin et al., 2011).

HBCDs were first observed in 1998 by Sellstrom et al. in fish and sediment samples from the river Viskan in Sweden and, since then, widely found in the environment and biota. The compound was detected in air samples of northern Sweden and Finland and in various animal samples including fish, seals, polar bears, glaucous gulls, and peregrine falcons from Eastern Greenland and Svalbard (de Wit et al., 2004; Verreault et al., 2005; Vorkamp et al., 2005; Jenssen et al., 2004). From those studies it can be concluded that HBCDs undergo long-range transport.

The most frequent HBCD exposure sources for humans are food, dust, air and consumer products (Covaci et al., 2006; Harrad et al., 2010; Goosey et al., 2010). There are two major pathways of compound exposure – dermal and oral as a result of inhalation of vapour and particles (ECRA, 2008). The main risk group is firstly people working at plants producing EPS (expanded polysterene) with HBCD; they are exposed to HBCD mostly via direct dermal exposure or inhalation. Measured levels of the pollutant in blood serum of industrial workers were up to 856 ng/g^{-1} (Thomsen et al., 2007), while average compound presence in other people was <1 ng/g^{-1} (AMAP, 2009) and accumulation through the environment or products is the major risk within this group (Marvin et al., 2011).

HBCD levels in dietary products are region-dependent (Shi et al., 2009; Roosens et al., 2009). The main source of exposure for humans is animal-origin food reach in fat, for instance, fish and meat (Shi et al., 2009). The highest concentration of the compound at the present moment detected in marine products, especially fish (Knutsen et al., 2008). Besides food, other important exposure route is through indoor air, in particular dust (Abdallah et al., 2008). HBCD contamination levels are especially high in toddlers as they ingest approximately 200 mg dust a day and this intake may be 10 times greater than dietary exposure level (Harrad et al., 2009).

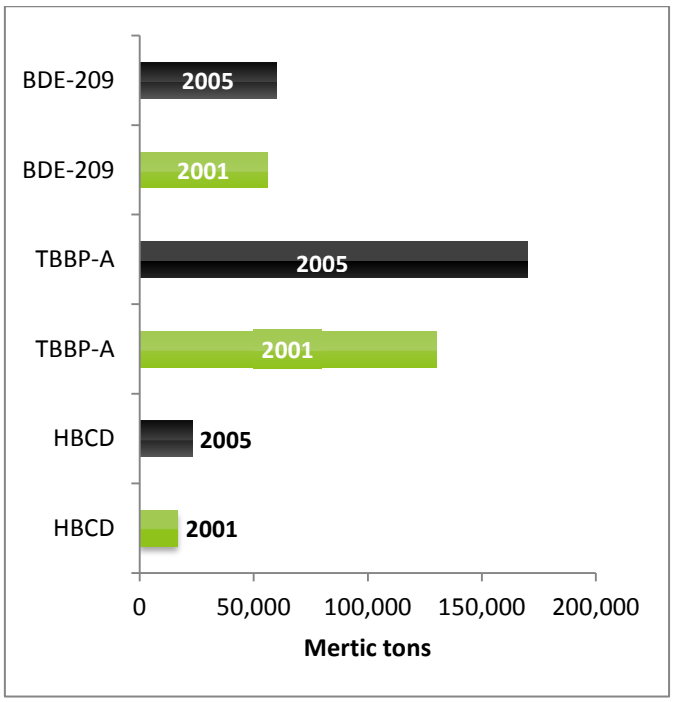


Chart 1.2. Consumption of various BFRs in 2001 and 2005, metric tons.

1.2.3. Toxicity

In contrast to polybrominated diphenyl ethers there are very few toxicological data available on hexabromocyclododecanes. Acute toxic effects of the compound are relatively low (Darnerud, 2003). However, there are some data concluding that oral exposure to HBCDs induces drug-metabolizing enzymes (for instance, hepatic cytochrome P450) in rats (Germer et al., 2006), and that HBCDs may cause genetic recombination leading to cancer (Helleday, 1999; Ronisz, 2004; Covaci et al., 2006). Moreover it is known that HBCDs affect thyroid function, brain development, neuron function, reproduction and development (Marvin et al., 2011). Furthermore, at the molecular level exposure to HBCD cause oxidative stress, decrease in antioxidant defense, disruption of calcium homeostasis, as well as induction of cell death (Fonnum & Mariussen, 2009).

Disruption of thyroid function proved to be a constant response to HBCD. Thyroid hyperplasia and lower circulating concentrations of thyroxine (T4) were reported in rats (ACC

1, 2001). Later other studies confirmed that HBCD induces thyroid toxicity both *in-vivo* and *in-vitro* in mammals, fish and birds (Marvin et al., 2011).

Furthermore, in studies on reproductive system in rats, high concentration of hexabromocyclododecanes suppressed oogenesis (Darnerud, 2003). *In-vivo* studies in rare minnow (*Gobiocypris rarus*) showed that HBCD induced apoptosis, increased formation of reactive oxygen species (ROS), causing oxidative damage to lipids, proteins, DNA and decline in antioxidant volumes (Zhang et al., 2008).

1.2.3.1.Neurotoxicity

HBCD influence on the mammalian nervous system is of a particular concern as neuronal cells are especially vulnerable to the toxicity. Several *in-vivo* studies conducted on HBCD observed neurotoxic effects of the compound (3–10 mg/kg) in rats. Developmental neurotoxic effects after neonatal exposure included aberrations in spontaneous behavior, learning and memory function (Eriksson, 2002). Elevation of the normal uptake in neurotransmitters in brain was also observed (Fonnum & Mariussen, 2009).

Very limited information is available regarding *in-vitro* studies of HBCDs neurotoxicity. Recent study showed that HBCD (1-30 µM concentrations) induce dose dependent neuronal cell death, ROS formation and increased [Ca^{2+}] levels, rapid depolarization of the mitochondria and cytochrome c release were also observed (Al-Mousa & Michelangeli, 2012).

1.3.TBBP-A

Tetrabromobisphenol A is a reactive flame retardant employed primarily in epoxy resins of printed circuit boards, but can also be used as an additive in acrylonitrile-butadiene-styrene (ABS) resins for electronic enclosures (Birnbaum & Staskal, 2004). Concentrations of TBBP-A in the environment and biota are quite low, most likely due to its main use as reactive FR (approximately 90%). TBBP-A is assumed to be less persistent than HBCD (as it is less toxic), and it (according to Covaci et al., 2009; however data not validated) presumably does not biomagnify and will unlikely be present at high concentrations in animal-origin food. In contrast no information is available for TBBP-A derivatives (Covaci et al., 2009).

1.3.1. Chemical and physical characteristics

Phenol, 4,4-isopropylidenbis 2,6-dibromo (TBBP-A) is a halogenated aromatic molecule (**Figure 1.1. b**) with a molecular mass of 543.87. TBBP-A, solid at room temperature is highly lipophilic and hydrophobic white crystalline powder. Commercial product is usually composed of 98% TBBP-A and other 2% made up of several brominated bisphenol A compounds (ACC 2, 2001). The industrial manufacturing process implies the bromination of bisphenol-A with a solvent addition, for instance, methanol, halocarbon, 50% hydrobromic acid or aqueous alkyl monoethers (EHC-172, 1995).

1.3.2. Occurrence and exposure

At the present moment TBBP-A is the most widespread brominated flame retardant in the world with a global consumption of 170,000 tons in 2004 (Birnbaum & Staskal, 2004). Tetrabromobisphenol-A is currently manufactured in the USA, Israel and Japan. In the EU it can be imported in different forms, e.g. as a primary product or as a complete or unfinished

products, for instance, plastics, printed circuit boards and electronic equipment. However, limited data are available on quantities imported in EU; overall amount seems to be approximately 40,000 tons per year (ERA, 2008). TBBP-A is also employed in derivative production, such as TBBP-A dibromopropyl ether, TBBP-A bis(allyl ether), TBBP-A bis(2-hydroxyethyl ether), TBBP-A brominated epoxy oligomer, and TBBP-A carbonate oligomers (EHC-172, 1995).

High concentrations of tetrabromobisphenol-A were detected in indoor air mostly nearby electronics dismantling plants. Furthermore, the compound was discovered in wastewater, dust, soil, sewage, sludge and sediment (Covaci et al., 2009). In spite of broad global use of TBBP-A, information on its presence in biota is deficient. The compound was found in eggs of Norwegian bird species (Herzke et al., 2005), in the blubber of porpoises (*Phocoena phocoena*) (Law et al., 2006) and in several aquatic organisms from the North Sea (Morris et al., 2004).

Generally, data regarding TBBP-A presence in human samples are scarce. Latest human studies report that TBBP-A was found mostly in plant workers, since they are occupied in the environment where exposure to the compound is much greater. It was also detected in 80% of the serum samples from computer technicians, reaching up to 3.4 pmol/g lipid (Jakobsson et al., 2002). Major exposure pathway for humans is inhalation of the compound, as its concentrations in food are rather low (Sjödin et al., 2003).

1.3.3. Toxicity

The available information on TBBP-A toxicity is limited. Acute oral toxicity seems to be low. However, there are data concluding that compound affects thyroid function, neuron function, reproduction and development. Moreover, TBBP-A can produce oxidative stress, decrease

antioxidant activity, disrupt calcium homeostasis and induce cell death (Al-Mousa & Michelangeli, 2012).

In contrast to PBDE and HBCD, TBBP-A does not accumulate in fatty tissues, but is bound to proteins and retained directly in blood. Due to its similarity to Bisphenol A - environmental estrogen – it is possible that TBBP-A may act as endocrine disruptor (Hughes et al., 2000). Moreover, 1–12 μM of TBBP-A is able to modify several cell signaling processes (Ogunbayo et al., 2007).

In several studies tetrabromobisphenol-A inhibited thyroid hormone transport function. It binds to human transthyretin (TTR) more effectively than thyroxin, showing 10 times higher potency over natural ligand (Meerts et al., 2000). TBBP-A also produced anti-thyroidal effect in Chinese hamster, as well as inhibited triiodothyronine (T3) binding to thyroid hormone receptors in ovary cells (Kitamura et al., 2005).

In fish, a concentration of 0.5 mg/l decreased survival and growth of young individuals. Studies with mice showed that TBBP-A cause decline in body weight, altered spleen weight, and decreased concentration of red blood cells, serum proteins, and serum triglycerides (IPCS, 1995). In a study with zebrafish (*Danio rerio*) exposed to water-borne TBBP-A, disorientation, lethargy, decrease in egg production and low reproductive activity were detected (Kuiper et al., 2007).

Even at low concentrations TBBP-A decreases fluidity of the cell membrane and therefore slowing down any biological processes associated with membranes (Lee et al., 1989). As the compound appears to be hydrophobic, it is possible that it will bind directly to phospholipid bilayers, and as a result it can be concluded that TBBP-A produces cytotoxic effect primarily on biological membranes (Pullen et al., 2003; Mariussen & Fonnum, 2002; Reistad et al., 2002).

1.3.3.1.Neurotoxicity

TBBP-A cause neurotoxic effects in cells, for instance, inhibits neurotransmitter uptake into synaptosomes and dopamine uptake into synaptic vesicles *in-vitro* (Pullen et al., 2003), what is followed by free radical formation (Reistad et al., 2002). At relatively low concentrations (1-30 µM) it can also raise the level of cytosolic Ca^{2+} in cerebellar granule cells, which triggers cell death mechanism (Reistad et al., 2007). Similarly to HBCD, TBBP-A was able to induce cell death, increase ROS formation and elevate [Ca^{2+}] levels in neuronal cell line, as well as cause depolarization of the mitochondria and cytochrome c release (Al-Mousa & Michelangeli, 2012).

1.4. BDE-209 (DBPE or decaBDE)

Polybrominated diphenyl ethers (PBDEs) – the third most commonly used class of brominated flame retardants – include 209 congeners, that are commercially classified as three major mixture groups: pentabrominated BDE, octabrominated BDE, and decabrominated BDE (Alaee et al., 2003). Decabrominated BDE or BDE-209 is employed in electronic equipment (Hardy, 2000) and represents more than 90% of all PBDE worldwide production (Fonnum & Mariussen, 2009). As all PBDEs BDE-209 is persistent organic pollutant, bioaccumulative in the environment and biomagnified by humans (Darnerud et al., 2003; Birnbaum & Staskal, 2004; Law et al., 2006).

1.4.1. Chemical and physical characteristics

BDE-209 (3,30,4,40,5,50,6,60-decabromodiphenyl ether) is a large, bulky molecule (**Figure 1.1. c**), with high molecular weight (959). DecaBDE mixture is made of >97% BDE-209, minor amounts of nonaBDEs (BDE-206, BDE-207, BDE-208), and octaBDEs (Goodman,

2009). This additive, fully brominated PBDE is almost insoluble in water and partially soluble in organic solvents (Costa & Giordano, 2011). DBPE is lipophilic, but due to its size the absorption through intestinal tract by passive diffusion is most likely impossible (Morck et al., 2003). However, recent studies showed that absorption may be solvent-dependent and increase from just 10 to 26% (Sandholm et al. 2003). From available data it is assumed that BDE-209 has a short half-life and is rapidly excreted, following little accumulation in tissues (Costa & Giordano, 2011).

1.4.2. Occurrence and exposure

PBDEs are extensively employed in a range of consumer products, for instance, textiles, carpets, polyurethane foams, electronic cables, television sets and computers. Despite the fact that pentaBDE and octaBDE were banned in the European Union and in the USA, BDE-209 is still globally used and manufactured in the United States and Europe (Costa & Giordano, 2007).

Due to additive nature of PBDEs there is a high probability that they can leach from the products into the environment. Following recent studies BDE-209 was detected in dust, air, sediments, soil, sludge and in various animal species (deWit, 2002; Hites et al., 2004; Law et al., 2006; Chen & Hale, 2010; USEPA, 2010). It was also detected in humans, mainly plant workers (Gill et al., 2004; McDonald, 2005) and in consumer products. For example, in Texas BDE-209 was found in 31 diverse food types (with a highest level in butter) (Schecter et al., 2010).

Exposure to decaBDE mainly occurs through food, domestic dust, and the occupational setting or can be a mix of all three pathways and is typically age-dependent (Johnson-Restrepo & Kannan, 2009).

1.4.3. Toxicity

BDE-209 cannot penetrate the cell wall as easy as HBCD or TBBP-A, due to its bulky configuration and low absorption rate. Acute and chronic toxicities of the compound are rather low, mainly targeting liver and thyroid gland. Furthermore, in some studies altered carcinogenicity (elevated tumor formation) was reported (Costa & Giordano, 2011). It is assumed that DecaBDE is not genotoxic (Hardy et al., 2009), and NOEL (No Observable Effect Level) values for it are relatively high – from a few hundred to several thousand mg/kg a day. However, recent studies reported that BDE-209 affects thyroid function, neuronal function and causes developmental effects. Moreover, TBBP-A can produce oxidative stress, decrease antioxidant activity, disrupt calcium homeostasis and induce cell death.

Various effects of BDE-209 were detected *in-vivo* (subacute/subchronic) studies in rats. At the dose of 80 mg/kg body wt. thyroid hyperplasia, liver enlargement and hyalin degeneration in kidney were seen (IPCS, 1994). Much less data available on BDE-209 effects *in-vitro* in comparison to lower brominated PBDEs. It is know that compound affects thyroid function disrupting thyroid hormone mediated transcription at very low concentration (0.01 nM) and inhibits thyroid hormone-induced dendrite arborization (Costa & Giordano, 2011).

1.4.3.1.Neurotoxicity

Limited information is available on neurotoxic effects of BDE-209. Several animal studies showed that the compound induces developmental neurotoxicity resulting long-lasting changes in locomotor activity and cognitive behaviour. A couple of studies also indicated that decaBDE directly affects neuronal cells (Costa & Giordano, 2011).

Experiments with rat neuronal cells evidenced that BDE-209 (dissolved in dimethylsulfoxide) has an effect on hippocampal neurons. At the concentrations of 10.4, 31.2 and 52 µM

(approximately 10, 30 and 50 µg/ml) decaBDE induced concentration-dependent reduction of cell viability, increased apoptotic cell death and elevated phosphorylation. It also caused oxidative stress, resulting ROS formation, malonyldialdehyde (MDA) and nitric oxide increase, as well as disrupted calcium homeostasis and reduced antioxidant (superoxide dismutase) capacity (Chen et al., 2010).

Another study also pointed out that at the same concentrations (10, 30 and 50 µM) of the compound declined rat hippocampal neuronal cell viability, increased apoptotic cell death and elevated intracellular calcium and ROS levels (Zhang et al., 2010).

Ultimately, in the most recent study Pellacani et al. (2012) reported that BDE-209 caused concentration-dependent DNA damage. In this experiment human neuroblastoma cells were treated for 4 and 24 hours with the concentrations of 5, 10, and 20 µM. In 4 hours significant DNA damage was observed, whereas after 24 hours it was considerably reduced, suggesting that DNA repair mechanism has been activated. However, following the treatment with the highest concentration no repair was detected (Pellacani et al., 2012).

1.5.Oxidative stress

1.5.1. Mechanism and specifics

Aerobic cells constantly consume oxygen, 90% of which is later utilised in electron transport chain as well as in nutrient oxidation, followed by the formation of carbon dioxide, water and energy. While about 5% of oxygen (under normal conditions) undergo partial one-electron reduction resulting addition of electrons and that leads to reactive oxygen species (ROS) formation (Lushchak & Semchyshyn, 2012).

Normally, ROS are rapidly removed from the cell by antioxidant defense system, for instance, superoxide dismutase (SOD) or glutathione (GSH) peroxidase (Tocher et al., 2002). However,

in a state where the ROS formation exceeds the antioxidant defense an imbalance between the ROS creation and elimination is formed. This condition is defined as oxidative stress (Zhang et al., 2008).

There are 3 main causes of oxidative stress: a) a rise in oxidant generation, b) a decline in antioxidant defence, c) a failure to repair oxidative damage. The consequences of oxidative stress are damage to proteins, lipids, DNA and a decline in antioxidant protection. Furthermore, oxidative stress is also associated with mitochondrial damage, because mitochondria are both employed in reactive oxygen species formation and serve as targets for ROS (Murphy, 2009). The defects formed in DNA are base modifications, single- and double-strand breaks and apurinic/apyrimidinic site creation, and all those may cause cancer if not repaired (Girard and Boiteux, 1997).

Oxidative stress may also result in the non-specific post-translational protein transformation and can induce aggregation of proteins. Human brain uses about 20% of inhaled oxygen, 90% of which is employed in production of energy during oxidative phosphorylation; therefore neurons are especially sensitive to oxidative stress. Neuronal cells are susceptible to oxidative damage due to increased metabolic activity, reduced antioxidant volumes and non-replicative nature (Lee et al., 2012).

1.5.2. Reactive oxygen species

Reactive oxygen species are extremely reactive due to unpaired valence shell electrons. They are created as a natural by-product of an oxygen cycle, but can also be produced by exogenous sources, for instance, ionizing radiation (Siddique, 2008). ROS are formed during the electron transport of mitochondria in a process called aerobic respiration or by

oxidoreductase enzymes and metal catalyzed oxidation. The most common reactive oxygen species are superoxide, hydrogen peroxide, hydroxyl radical and nitric oxide (Held, 2010).

For a long time it was clear that ROS are involved in the immune response to microbial invasion, eliminating bacteria in cells. However, only recent studies provided evidence that they play a crucial role as a messenger in signal transduction and cell cycle, including apoptosis and gene expression (Hancock et al., 2001). Elevated ROS levels can induce either survival or apoptotic death of the cell, what depends on exposure time and concentration. ROS are in charge of such processes as regulation of the blood pressure and redox signaling-mediated enzyme regulation (Hou et al., 1999). Nevertheless, high levels of reactive oxygen species in response to, for example, UV or heat exposure, are toxic for aerobic organisms and lead to oxidative stress (Siddique, 2008).

1.5.3. Antioxidant system. Glutathione.

Antioxidant system is a range of mechanisms that protects cells from the destructive effects of ROS. This system includes a number of enzymes and non-enzymatic molecules such as superoxide dismutase, catalase, glutathione peroxidase and vitamin C. Glutathione (GSH) plays one of the most important roles in protection of cells against ROS (Held, 2010).

GSH is an ubiquitous thiolic tripeptide (glutamyl-cysteinyl-glycine) that contains a sulphydryl group (Jefferies et al., 2003). This group serves as a target for reactive molecules, and as a result glutathione is converted from the reduced (GSH) to oxidized form (Jones, 2000). Furthermore, reacting with other reactive glutathione (and forming GSSG) this thiol is rapidly converted back to GSH. Various functions of GSH include modulating of cell response to redox changes, detoxification of the drug metabolites, gene expression control and apoptosis

regulation, as well as involvement in transmembrane transport of organic solutes (Jefferies et al., 2003).

1.6.Cell death

PCD (Programmed cell death) is a type of cell death encoded in genes and is crucial for cell development and homeostasis of organisms. In embryonic development 3 main mechanisms of programmed cell death can be distinguished: apoptosis (cell shrinkage and chromatin condensation), necrosis (not always programmed; damage of plasma membrane and intracellular content spillage) and autophagy (autophagic vacuole formation). Which mechanism will be triggered strictly depends on the nature of the signal, developmental stage and the cell type as all mechanisms are connected and one death program can be inhibited and replaced by the other (Clarke, 2002).

1.6.1. Apoptosis

Apoptotic cell death is a main route employed in elimination of unnecessary, old or damaged cells, essential for embryonic development, tissue homeostasis and immune regulation (Ellis et al., 1991). It is a carefully controlled process with specific morphological and biochemical mechanisms. Apoptosis is induced by physiologic and pathologic stimuli and for the normal functioning it requires signaling cascade activation, which is regulated by caspases. During the apoptotic death the following chain of processes occurs: margination and fragmentation of chromatin, cellular shrinkage and fragmentation, and consumption by surrounding cells (Kerr et al., 1972). It is believed that apoptotic mechanism is a consequence of a caspase-controlled cascade and any divergence from it will cause the cell death by a necrotic pathway (Oppenheim et al., 2001).

In the cells that undergo apoptosis decreased levels of adenosine triphosphate (ATP) transforming to ADP are detected, what may happen due to reduced synthesis of ATP in the mitochondria (Eguchi et al., 1999). Also intensive peroxidation is observed in lipid bilayers, what result in modification of cell membrane composition.

Apoptosis may occur by two different but interconnected routes – the cell surface death receptor-induced (extrinsic) and the mitochondria-induced (intrinsic) mechanism. The main components of the extrinsic mechanism are the adapter protein FADD (Fas-Associated Death Domain protein), the death-regulating protease enzyme caspase-8, and FLIP that controls caspase-8 activity. These apoptotic proteins are also employed in embryonic development (Zeiss, 2003).

1.6.2. Necrosis

For a long time the nature of necrosis was regarded as a result of accidental and uncontrolled physico-chemical stress. However, recently it was proved that necrotic death like apoptosis is highly controlled and is pathologically and physiologically significant death pathway, as it can eliminate oncogenic cells, which have an ability to avoid apoptosis (Festjens et al., 2006). Necrosis can be defined as a process where unrepairable damage of plasma membrane, swelling of cytoplasm and collapse of organelles occurs (Fiers et al., 1999). The degradation of DNA in necrotic cells proceeds randomly involving a smear-mechanism (Higuchi, 2003).

Several studies showed that serine/threonine kinase RIP1, which includes a death domain, plays a central role in the initiation of necrosis. Major factors that induce necrotic processes are disruption of calcium homeostasis and reactive oxygen species formation. They directly or indirectly cause damage to proteins, lipids and DNA resulting failure of organelles and cell membrane breakdown. Necrotic cells are then lysed, provoking pro-inflammatory signaling

cascades (cytokine release) and forcing the contents to leak into the extracellular space (Festjens et al., 2006).

1.6.3. Autophagy

Autophagic cell death is a process where lysosome-mediated degradation of unwanted or damaged cellular elements occurs. On a physiological level it is used to keep the organelle biogenesis, protein synthesis and their elimination in balance. It also serves as a significant mediator of pathological responses and is along with ROS and RNS (reactive nitrogen species) employed in cell signalling processes as well as in damage of the proteins (Lee et al., 2012).

There are three classes of autophagy: macroautophagy (classic autophagy), microautophagy and chaperone-mediated. Macroautophagic mechanism involves the creation of double-membrane vesicles called autophagosomes that grow and merge with lysosomes for their content degradation using acidic hydrolases (Schneider & Zhang, 2010). In microautophagy, on the other side, cytosolic components are directly enclosed in lysosomes and ingested by them employing membrane involution (Mijaljica et al., 2011; Sahu et al., 2011). Whereas, chaperone-mediated autophagy utilizes chaperones to target proteins similar to the pentapeptide KFERQ, transporting them to the lysosomes for the elimination (Bejarano & Cuervo, 2010).

It is crucial for the cells to control autophagic process at all stages, especially clearance of damaged proteins generating excessive ROS and removal of whole organelles by mitophagy. If the process is disrupted at early stage, that results in accumulation of ubiquitous proteins, elevated ROS and failure of mitochondria. Likewise, reduced removal of dysfunctional

proteins, organelles and DNA via autophagy cause mutations and leads to tumorigenesis (Chen & Karantza-Wadsworth, 2009).

1.7.DNA damage

Initiation of DNA damage is one of the primary stages of chemically caused oncogenesis. There are two main pathways of DNA lesion: direct attack of the chemical or its metabolites and indirect mechanism of damage related to oxidative stress (Kong et al., 2011). Brominated flame retardants are believed to cause DNA damage through ROS formation, as they do not result in gene mutations (Pellacani et al., 2012). The consequences of the oxidative stress-related mechanism are single or double strand breaks and the formation of oxidation products such as abasic sites and oxidized bases. DNA single-strand breaks are the most common oxidative damage in cells. Their repair is especially important for neurons, as failure to fix the damage can lead to mutagenic processes in brain (Hwang & Kim, 2007; Van Loon et al., 2010; Kong et al., 2011).

1.8.Aim and objectives

The aim of the study is to investigate molecular mechanisms of neurotoxicity by brominated flame retardants in SH-SY5Y human neuroblastoma cells.

Objectives:

- To evaluate at which concentrations BFRs influence cell viability and determine non-toxic concentrations;
- To measure levels of oxidative stress (ROS assay and GSH assay);
- To assess levels of DNA damage using Alkaline Comet assay;
- To obtain data regarding the prevailing mechanism of cell death (Apoptosis assay, Autophagy assay).

2. Materials and Methods

2.1. Chemicals

All chemicals were purchased from Sigma-Aldrich unless stated otherwise.

- **BDE-209 –** (97% pure) 1.4 mM solution dissolved in DMSO
- **DMEM –** Dulbecco's modified Eagle's medium with 4500 mg/L glucose, L-glutamine and sodium pyruvate, pyridoxine, endotoxin tested, sterile filtered. Contains 2 mM L-glutamine, 1% penicillin (100 units/ml), streptomycin (100 µg/ml), 10% (vol/vol) heat-inactivated foetal bovine serum (FBS) and supplemented with 1% (vol/vol) non-essential amino acid solution (MEM).
- **DMSO** – Dimhethyl sulfoxide (Fisher scientific), sterile
- **HBCD –** (95% pure) 2 mM solution dissolved in DMSO
- **PBS –** Phosphate Buffered Saline (Dulbecco A) (Oxoid, Basingstoke), containing Sodium Chloride (8.0), Potassium chloride (0.2), Di-sodium hydrogen phosphate (1.15), potassium dihydrogen phosphate (0.2).
- **TBBP-A –** (97% pure) (Acros Organics, UK) 4 mM solution dissolved in DMSO

2.2. Cell culture

SH-SY5Y neuroblastoma cells (a gift from Frank Michelangeli) were cultured in sterile conditions in a Bio Air Aura B4 Class II hood. The hood was sterilized with 70% ethanol spray every time before and after use. The working instruments were also disinfected with ethanol before application.

2.2.1. Defrosting cells

Cell vial (with 1 ml solution containing frozen cells) was extracted from liquid nitrogen and placed in a 37 ºC water bath. When unfrozen, the vial was immediately transferred to the hood and all contents were added into 25cm^2 cell culture flask (T25) (BD Falcon™), containing 5 ml of warm DMEM media. The flask was then placed into a humidified chamber (5% CO2, 95% air; MCO-15AC, Sanyo, Japan) and incubated at 37 °C.

2.2.2. Cell sub-culture

When the cells were confluent enough (occupying about 80% of the flask), they were sub-cultured into 75cm^2 cell culture flask (T75) (BD Falcon™). Firstly, DMEM was removed from T25 and the cells were washed with sterile PBS. After removing PBS trypsin-EDTA (2 ml) was added to break the attachments between cells, and the flask was incubated in humidified chamber at 37 °C for approximately 2 minutes. DMEM media (8 ml) was added to the flask and the solution was transferred into 15 ml Falcon tube (BD Falcon™) and centrifuged at 1500 rpm for 5 minutes (MSE Falcon 6/300, Sanyo, Japan). Supernatant was then removed and cell pellet was resuspended in 1 ml of media, and added to T75 flask with 15 ml of fresh DMEM. During the study cells were sub-cultured and maintained in T75 flasks.

2.2.3. Cryopreservation

Cells were trypsinised and pelleted in a centrifuge (Falcon 6/300) as described above. After supernatant removal cell pellet was resuspended in 2 ml of freezing media (1:9 mixture of sterile DMSO (10% v/v) and Feotal Bovine Serum) and this solution was transferred to 2

cryovials (1 ml each). Cryovials stayed in -80 °C freezer overnight and then were transferred into liquid nitrogen.

2.3. Compound dilution

The right concentrations of compounds were reached using serial dilution. All concentrations used for HBCD, TBBP-A and BDE-209 are summarized in **Chart 2.1**. Firstly, media for dilution was prepared (DMEM + 1% DMSO) to obtain same concentration of DMSO in all samples. Then the highest compound concentration was reached adding the required compound's stock solution (stored in a fridge) and DMEM (1:10 in case of BDE-209 or 1:100 with HBCD, TBBP-A) into 15 ml Falcon tube. The next concentration was reached taking the solution from the first tube and adding the same amount of media with DMSO (1:1). This dilution was continued till the last tube (except TBBP-A concentration of 15 µM, where dilution coefficient was 1:1.67) and finished reaching the lowest concentration.

Compound	Concentration, µM							
HBCD	20	10	5	2.5	1.25	0.625	0.313	0.156
TBBP-A	40	20	15	10	5	2.5	1.25	0.625
BDE-209	240	120	60	30	15	7.5	3.75	1.875

Chart 2.1. BFR concentration spectra (µM) used for MTT assay.

2.4. Viability (MTT) assay

Cell viability assay was performed (according to Mosmann, 1983) to estimate levels of BFR cytotoxity. Cells were cultured in 96-well plates (Corning). The media (100µl) was removed and fresh media containing all range of compound concentrations (diluted as above, including 0 concentration – control) was added. Cells then were incubated for 24 hours at 37 ºC.

MTT (3-[4,5-Dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide) stock solution in PBS (5mg/ml) was prepared. Compound containing media was subsequently replaced with MTT stock solution in DMEM (final concentration 0.5 mg/ml) and stored in humidified chamber (37 ºC) for 2 hours. After MTT media was removed and DMSO was added to each well (and to 3 spare wells as blank). The absorbance was read employing Bio-Tek FL600 microplate reader (Bio-Tek Instruments Inc. USA) at 590 nm.

2.5. Fluorescence spectroscopy

Levels of oxidative stress (Reactive Oxygen Species formation and glutathione) were measured using Fluorescein ROS assay and GSH assay.

2.5.1. Fluorescein ROS assay

Firstly, cells were cultured in 6-well plates (Corning) and incubated overnight at 37 ºC. Then non-toxic compound concentrations (below LC50 – lethal concentration 50%) were added and plates were incubated in humidified chamber at 37 ºC for 24 hours. The concentrations used for ROS assay are displayed in **Chart 2.2**. After 23 hours 25 mM H_2O_2 (hydrogen peroxide) was added to one well (for positive control) for 1 hour. Then compound-containing media was replaced by DCFDA (2',7' –dichlorofluorescein diacetate, Cellular Reactive Oxygen Species Detection Assay Kit) mixture with DMEM (100 µM final concentration) and plates were incubated for 60 minutes at 37 ºC. Again media was removed and cells were washed with PBS. Sodium hydroxide (10 mM) was subsequently added, cells were scraped using the cell scraper and incubated at room temperature for 15 minutes. After cells were transferred to eppendorfs, centrifuged at 13 000 rpm for 5 minutes using a bench-top centrifuge (MSE Micro Centaur, Sanyo) and supernatant was then transferred to 3 ml cuvettes. The

fluorescence was read using Luminescence spectrometer (Perkin Elmer LS50B, UK; excitation wavelength 490 nm and emission wavelength 520 nm).

Compound	Concentration, µM			
HBCD	5	2.5	1.25	0.625
TBBP-A	20	10	5	2.5
BDE-209	15	7.5	3.75	1.875

Chart 2.2. BFR concentrations (µM) used for Fluorescence spectroscopy and Comet assay.

2.5.2. GSH assay

Buffers:

- ✓ **Cell lysis buffer:** 0.1% Triton X-100 in PO_4-EDTA assay buffer
- ✓ **Protein precipitation buffer (TCA 50%):** TCA (Trichloroacetic Acid) diluted in dH_2O (0,5 g/ml)
- ✓ **PO_4-EDTA assay buffer:** 100 mM NaH_2PO_4, 5 mM Na_2EDTA in dH_2O; pH – 8.0.
- ✓ **O-Phthalaldehyde (OPT):** OPT powder in 100% Methanol (1 mg/ml)

Cells were cultured in 6-well plates and treated with compounds as in ROS assay. After 24 hours media was removed and cells were washed in PBS. Then ice-cold cell lysis buffer was added, cells were scraped into eppendorfs and placed on ice. TCA 50% (vol/vol) was subsequently added to lysate, which was centrifuged using a benchtop microfuge at 13 000 rpm for 5 minutes. After that GSH containing supernatant was transferred to a fresh eppendorf. At this stage supernatant was either placed in -80 °C freezer or the assay was immediately continued.

GSH stock was prepared: warmed GSH was mixed with PO_4-EDTA assay buffer (0.1 mg/ml). Fluorescence cuvettes (3 ml) were used for the assay. PO_4-EDTA was added to all cuvettes and GSH was transferred to GSH-cuvettes for a standard curve of 0, 1, 2, 4, 6, 8, 10, 12, 14,

16, 18 and 20 µg. All GSH containing cuvettes were then supplemented with TCA 5% (diluted in water 1:10). After supernatant from each sample was added to the sample-cuvettes. Finally, OPT was added to every cuvette, the cuvette-box was shacked, covered and left at room temperature for 15 minutes. The fluorescence was read at excitation wavelength of 350 nm and emission wavelength of 420 nm.

2.5.3. Bradford assay

After GSH and ROS assays protein concentrations in all samples were determined using Bradford assay (Bradford, 1976). First, Bio-Rad protein assay solution was diluted in destiled water (1:5), filtered and added to all 1 ml cuvettes. Then BSA (protein) was mixed with PBS in eppendorf and added to cuvettes for a standard curve of 0, 2, 4, 6, 8 and 10 µg. Cell samples were added into the sample cuvettes and mixed. The absorbance was read at 595 nm.

2.6. Alkaline comet assay

Buffers:

- ✓ **Lysis Buffer:** 2.5 M NaCl, 0.1 M Na_2EDTA, 10 mM Tris Base dissolved in H_2O, pH adjusted to 10.0 and supplemented with sodium lauryl sarcosinate (stored at 4°C). Before use an aliquot of the buffer was mixed with Triton X-100 (1% final concentration) and DMSO (10% final concentration)
- ✓ **Electrophoresis Buffer:** 9 M NaOH, 20 mM Na_2EDTA in dH_2O (stored at 4 °C)
- ✓ **Neutralization Buffer:** 0.4 M Tris Base dissolved in H_2O, pH 7.5 (stored at 4°C)

Slides for the comet assay (Singh et al., 1988) were prepared at least 2 days before: each slide was covered with NMPA (Normal melting point agarose 0.5% w/v in PBS) from 1 side. Cells

were cultured and treated (same concentrations as for ROS assay) in 6-well plates. After 23 hours 50, 25 or 12.5 mM H_2O_2 was added to 1 well (for a positive control) for 1 hour. Media was removed and cells were washed in PBS. Then cells were either scraped in PBS or trypsinised (modified protocol), added into eppendorfs and pelleted down in the benchtop centrifuge at 8 000 rpm for 5 minutes. The pellet was resuspended in PBS. LMPA (Low melting point agarose 0.5% w/v in PBS) was then warmed up, placed into eppendorf and mixed with sample. The mixture was used to coat a slide and a coverslip was added on top. Slides were left on ice for approximately 20 minutes to solidify. When solid, the coverslip was removed and slides were transferred to a Coplin jar containing lysis buffer. The jar was wrapped in foil and stored in the cold room for 1 hour. After slides were removed from the Coplin jar and placed into a large electrophoresis tank, containing electrophoresis buffer, for 20 minutes. A power pack was turned on at 32 V (300 mA) and electrophoresis was carried out for 20 minutes. Then slides were removed and washed with neutralisation buffer 3 times for 5 minutes. SYBR gold solution was subsequently prepared (SYBR gold in neutralisation buffer, 1:1000), added to each slide and the coverslip was placed on top. Finally, slides were transferred to moist box, covered with foil and stored in the cold room overnight. DNA damage was evaluated in each slide (100 cells counted) under Axiovert 10 Microscope (Zeiss, West Germany) using AVT Smartview camera and Comet 4 Software.

2.7. Vybrant apoptosis assay

Buffers:

- ✓ **1x Annexin-Binding Buffer:** 50 mM HEPES, 700 mM NaCl, 12.5mM $CaCl_2$, pH 7.4. 5x Annexin-Binding Buffer dissolved in H_2O (1:4)
- ✓ **FITC annexin V:** 25 mM HEPES, 140 mM NaCl, 1 mM EDTA, pH 7.4, with 0.1% bovine serum albumin (BSA)

- ✓ **Propidium iodide (100 µg/ml):** PI solution in 1x Annexin-Binding Buffer

Cells were cultured and treated (HBCD – 5 µM, TBBP-A – 10 µM and BDE-209 - 15µM in supplemented or clear unsupplemented DMEM) in 6-well plates. Then media was removed and cells were washed in PBS. Trypsin was added, and cells were incubated for 10 minutes at 37 °C. Again PBS was added, cells were transferred to eppendorfs and centrifuged at 8 000 rpm for 5 minutes. Supernatant was removed, pellet was resuspended in PBS and re-centrifuged. After cells were resuspended in 1x Annexin-Binding Buffer. Propidium iodide and FITC annexin V were added to the suspension, and cells were incubated for 15 minutes at room temperature. After the incubation period, 1x Annexin-Binding Buffer was added, samples were mixed and transferred to the Round-bottom tubes (Falcon®, 352053, Becton Dickinson, USA). Stained cells were analyzed using flow cytometer (BD FACS Calibur), measuring the fluorescence emission at 470 nm (FITC) and 530 nm (PI).

2.8. Transformation

First, LB broth (10 g/l Bacto-tryptone, 5 g/l Bacto-yeast extract, and 5 g/l NaCl, pH 7.5 in dH_2O) and LB/Agar (LB broth, technical agar in dH_2O) bacterial growth medium were prepared and sterilized by autoclaving. Then kanamycin antibiotic (30 µg/ml) was added to the LB/Agar medium and mixture was transferred to Petri plate (Falcon® 353003, Becton Dickinson, USA) and left to solidify.

E. coli Super Competent cells (JM109, Promega) were transformed using pEGFP-C1-LC3 vector (BD Biosciences). Cells were removed from -70 °C freezer, thawed and transferred to 15 ml tubes. Then pEGFP-C1-LC3 (50 ng) was added and tubes were placed on ice for 10 minutes. Cells were heat-shocked in water bath for 45-50 seconds (42 °C) and returned on ice for 2 minutes. LB broth (with added kanamycin, 30 µg/ml) was mixed with cells and tubes

were incubated for 1 hour at 37 °C (225 rpm) in rotator (Aerotron INFORS AG). After cells were transferred to Petri plates and incubated at 37 °C for 12 hours. Grown *E.coli* colonies were used for DNA plasmid isolation experiments.

2.9. Plasmid DNA isolation

DNA was isolated using ISOLATE Plasmid DNA Mini Kit (Bioline). First, selected colony was cultured (in 50 ml tubes, containing LB with kanamycin, 30 µg/ml) in rotator overnight. Then the tubes were centrifuged at 6000 rpm for 10 minutes (Mistral 2000, Meadowrose Scientific LTD) and supernatant was resuspended in Resuspension buffer. Lysis buffer was added and suspension was mixed by inverting. Next, Neutralization buffer was added to the mixture and it was centrifuged for 15 minutes at 13 000 rpm to pellet cell debris. After the sample was transferred to Spin Column P, inserted in Collection tube and centrifuged for 1 minute at 10 000 rpm. Filtrate was removed, Wash buffer AP was added and tubes were spinned as above. Then Wash buffer BP was added and samples were centrifuged again, following the filtrate removal and re-centrifugation at 13 000 rpm for 2 minutes. Spin Column P was placed into Elution tube and Elution buffer was added. Samples were incubated at room temperature for 1 minute and centrifuged at 10 000 rpm.

2.10. Transfection

Cells were cultured and transfected in 6-well plates employing Thermo Scientific TurboFect Transfection reagent. Serum-free media (unsupplemented DMEM) was mixed with isolated plasmid DNA (4 µg) and TurboFect reagent (6 µl). Mixture was incubated at room temperature for 15 minutes and then added to each well. For the efficient transfection cells

were incubated at 37 °C for 48 hours. GFP (green fluorescence protein) expression was monitored using Axiovert 10 Microscope.

2.11. Cell fixing and Autophagy assay

After transfection cells were treated with compounds (HBCD – 5 µM, TBBP-A – 10 µM and BDE-209 - 15µM) and fixed after 0, 4 and 24 hours. Media was removed and cells were washed in PBS. Paraformaldehyde solution, 4% (paraformaldehyde powder, PBS, pH 7.4) was added to the cells, following the incubation for 15 minutes at room temperature. Then cells were washed in PBS (3 times) and left in a fridge (wrapped in foil) until required. Autophagic puncta was monitored employing fluorescent Nikon microscope (Japan).

3. Results

To investigate neurotoxic effects of brominated flame retardants in human neuroblastoma cells, they were treated with diverse concentrations of HBCD, TBBP-A and BDE-209. Cell viability levels and non-toxic concentrations (lower than LC50 – lethal concentration 50%) of the compounds were determined using MTT assay. These concentrations were further tested to evaluate levels of ROS and GSH – main oxidative stress indicators employing Fluorescence spectroscopy and DNA damage using Comet assay. All assays were repeated at least 3 times. In addition, the preliminary results on the mechanism of cell death were analyzed by apoptosis and autophagy assays.

3.1. MTT assay

An impact of brominated flame retardants on cell viability was measured to estimate the levels of their cytotoxicity. Wide range of concentrations (**Chart 2.1**) was used to evaluate the survival of cells after BFR exposure. Colorimetric MTT assay, where MTT agent is reduced to a violet formazan product (in metabolically active cells), was used to determine the number of viable cells.

All compounds showed concentration-dependent decrease in cell viability (**Chart 3.1-3.3**), but HBCD proved to be the most toxic for the cells. There was a slight shift in number of viable cells at very low concentrations with HBCD (from 0.156 – 1.25 µM) and TBBP-A (from 1.25 – 5 µM), which was followed by a sharp drop in viability at the concentrations 2.5 and 10 µM respectively. In contrast the third compound – BDE-209 constantly followed the trend of gradual reduction in cell viability.

LC50s were calculated to estimate at which concentration cell viability was 50% and LC50 values (or lower) were used for further experiments. The lethal concentrations, 50% were between 2.5 and 5 µM, 15 and 20 µM, 30 and 60 µM for HBCD, TBBP-A and BDE-209 respectively.

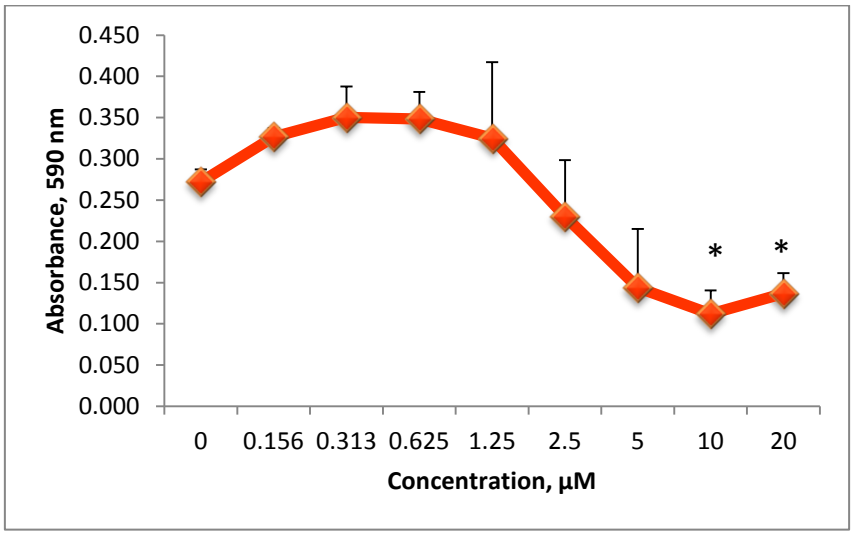


Chart 3.1. Influence of HBCD on SH-SY5Y cell viability. Cells were incubated with HBCD for 24 hours at different concentrations (0-20 µM). Cell viability (abs. 590 nm) was measured using MTT assay. Final absorbance calculated against DMSO blank, data points are mean values of 9 measurements ±Standard error (3 independent experiments). *Asterisks indicate significant values ($p<0.05$) calculated by Student's T-test.

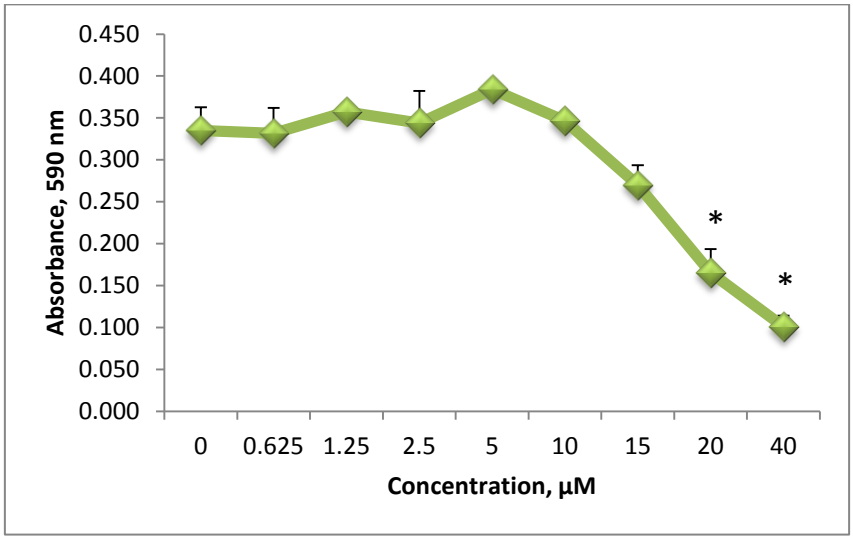


Chart 3.2. Influence of TBBP-A on cell viability. Cells were incubated with TBBP-A for 24 hours at different concentrations (0-40 µM). Cell viability (abs. 590 nm) was measured using MTT assay. Final

absorbance calculated against DMSO blank, data points are mean values of 9 measurements ±Standard error (3 independent experiments). *Asterisks indicate significant values ($p<0.05$) calculated by Student's T-test.

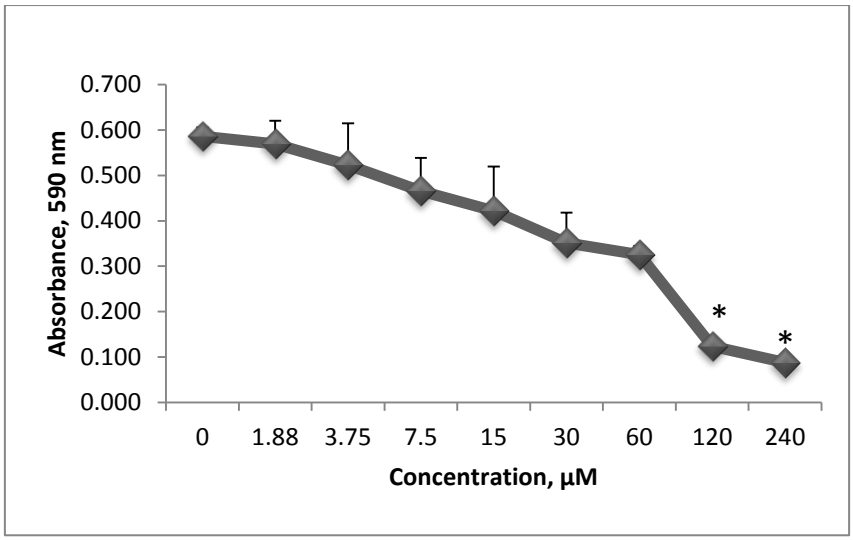


Chart 3.3. Influence of BDE-209 on cell viability. Cells were incubated with BDE-209 for 24 hours at different concentrations (0-240 µM). Cell viability (abs. 590 nm) was measured using MTT assay. Final absorbance calculated against DMSO blank, data points are mean values of 9 measurements ±Standard error (3 independent experiments). *Asterisks indicate significant values ($p<0.05$) calculated by Student's T-test.

3.2. Fluorescence spectroscopy

3.2.1. Fluorescein ROS assay

ROS formation was measured to determine the oxidative stress in cells. Four concentrations (**Chart 2.2**) for each compound were selected for the assay. Reactive oxygen species were detected using fluorescein ROS assay, where DCFDA reagent is oxidized by ROS into fluorescent DCF (2', 7' –dichlorofluorescin). The efficiency of the assay was verified by a positive control (50 µM H_2O_2), which cause elevated ROS formation. The fluorescence was measured between 490 and 520 nm and converted to "fluorescence per µg of protein", employing Bradford assay. During the experiments TBBP-A concentration of 20 µM proved

to be toxic as was killing the cells in 24 hours (before the assay) and therefore was excluded from the analysis. One of three compounds – HBCD (**Chart 3.4**) showed no concentration-dependent response in the assay. ROS formation with BDE-209 (**Chart 3.6**) slightly increased, but no in statistically significant manner. In contrast considerable concentration-dependent increase in ROS was detected with TBBP-A (increasing more than two times at the highest concentration – 10 µM in comparison to the control, **Chart 3.5**).

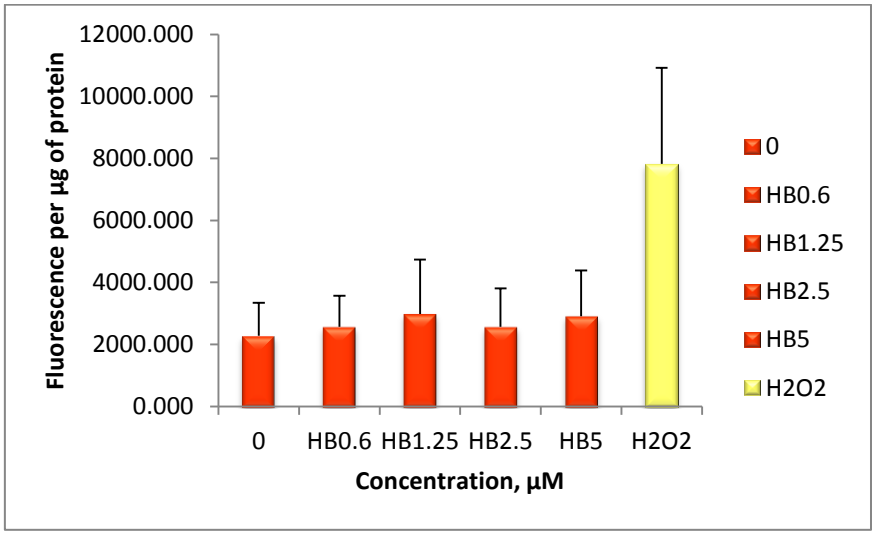


Chart 3.4. ROS production in HBCD treated SH-SY5Y cells. Fluorescein oxidation was detected using DCFDA after 24 hours treatment at different HBCD concentrations (0-5 µM) and compared to positive control (H_2O_2). Final values were calculated proportionally to the amount of protein (µg) obtained in Bradford assay. Data points represent mean values of 3 measurements ±Standard error (3 independent experiments).

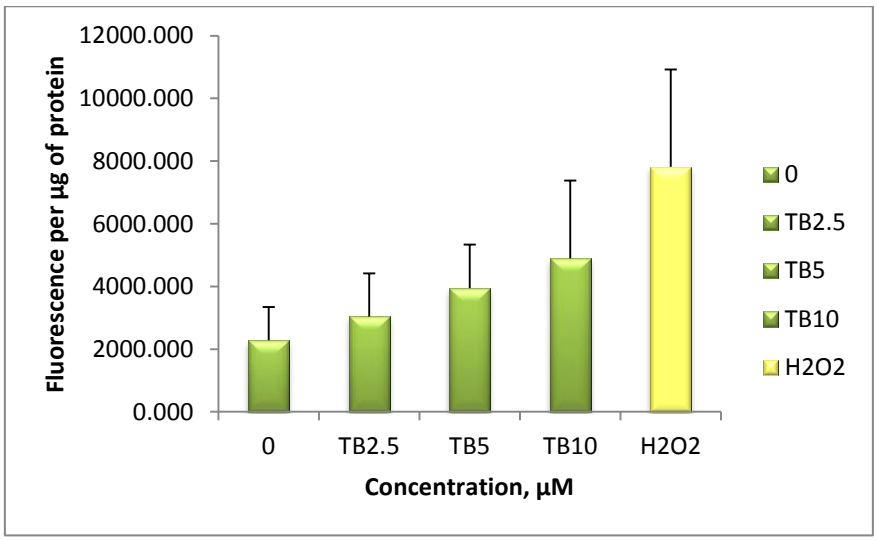


Chart 3.5. ROS production in TBBP-A treated cells. Fluorescein oxidation was detected using DCFDA after 24 hours treatment at different TBBP-A concentrations (0-10 µM) and compared to positive control (H_2O_2). Final values were calculated proportionally to the amount of protein (µg) obtained in Bradford assay. Data points represent mean values of 3 measurements ±Standard error (3 independent experiments).

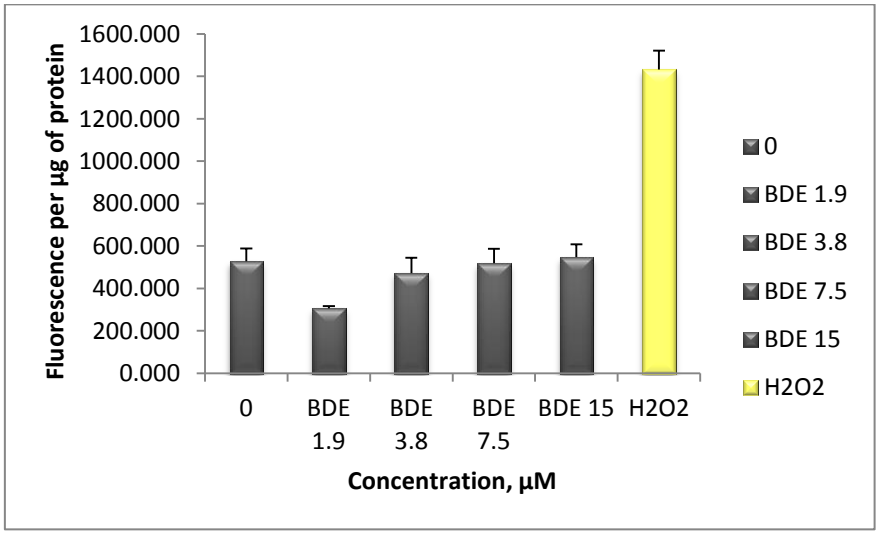


Chart 3.6. ROS production in BDE-209 treated cells. Fluorescein oxidation was detected using DCFDA after 24 hours treatment at different BDE-209 concentrations (0-15 µM) and compared to positive control (H_2O_2). Final values were calculated proportionally to the amount of protein (µg) obtained in Bradford assay. Data points represent mean values of 3 measurements ±Standard error (3 independent experiments).

3.2.2. GSH assay

In conjunction with the ROS assay the levels of glutathione were measured to indicate oxidative stress in cells. The concentrations of the compounds were selected as in ROS assay (also excluding 20 µM concentration). GSH assay was used to detect levels of glutathione in cells, involving transformation of luciferin derivative into fluorescent luciferin in the presence of GSH. The fluorescence was measured between 350 and 420 nm and was converted into µg of GSH per µg of protein (involving Bradford assay).

Two compounds – TBBP-A and BDE-209 (**Chart 3.8., 3.9**) showed only random insignificant changes in GSH levels, therefore no response was detected. On contrary, HBCD (**Chart 3.7**) considerably decreased levels of GSH. The results indicated more than 3 fold depletion of GSH at the highest concentration of 5 µM compared to the control.

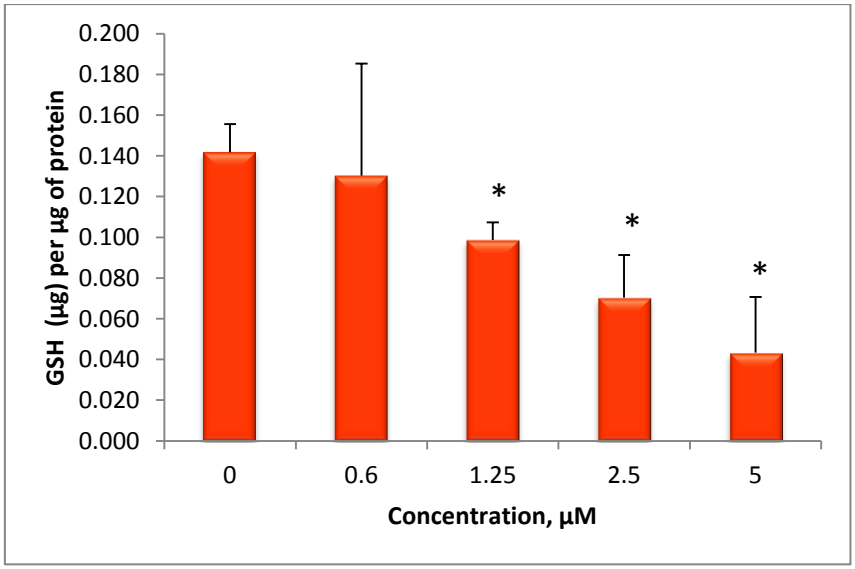


Chart 3.7. GSH quantities in SH-SY5Y cells treated with HBCD. GSH fluorescence was measured after 24 hours (0-5 µM concentration range) treatment with HBCD using GSH assay. Fluorescence was converted into µg of GSH present in 1 µg of protein (Bradford assay). Data points are mean values of 3 measurements ±Standard error (3 independent experiments). *Asterisks indicate significant values ($p<0.05$) calculated by Student's T-test.

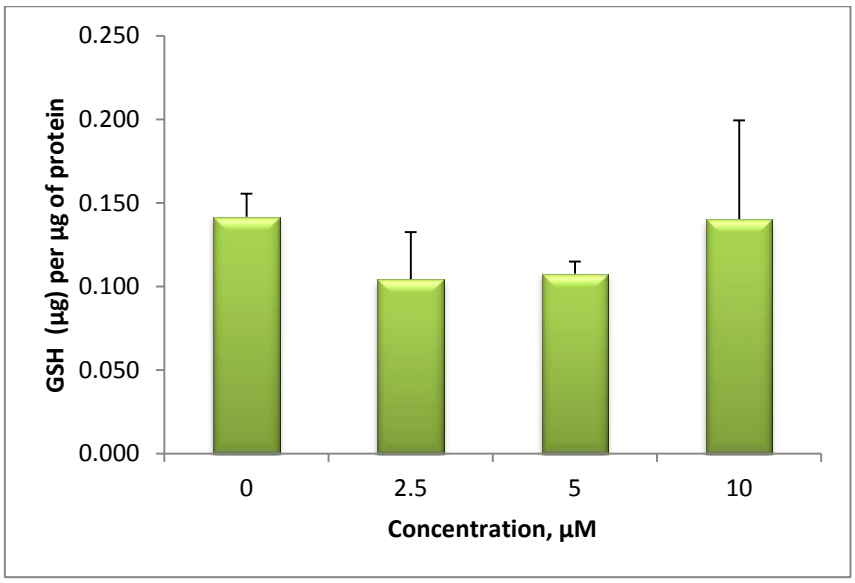


Chart 3.8. GSH quantities in cells treated with TBBP-A. GSH fluorescence was measured after 24 hours (0-10 µM concentration range) treatment with TBBP-A using GSH assay. Fluorescence was converted into µg of GSH present in 1 µg of protein (Bradford assay). Data points are mean values of 3 measurements ±Standard error (3 independent experiments).

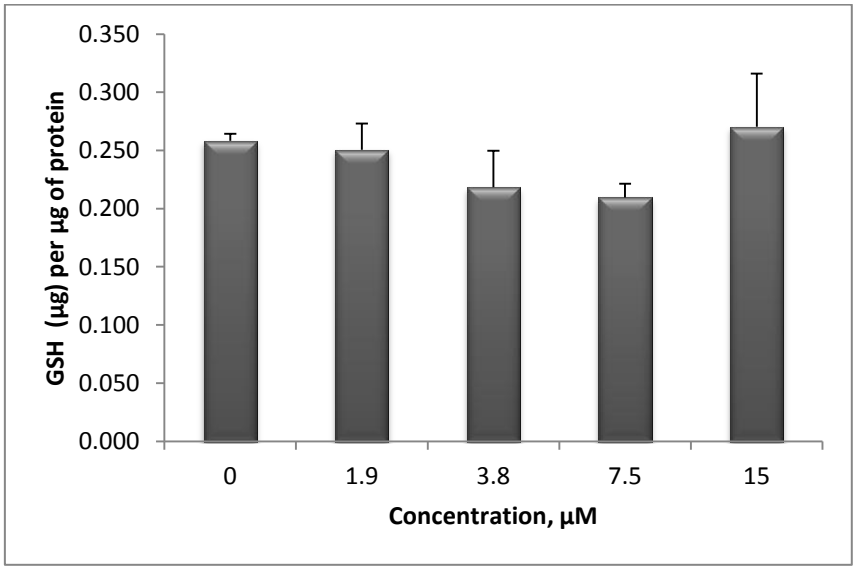


Chart 3.9. GSH quantities in cells treated with BDE-209. GSH fluorescence was measured after 24 hours (0-15 µM concentration range) treatment with BDE-209 using GSH assay. Fluorescence was converted into µg of GSH present in 1 µg of protein (Bradford assay). Data points are mean values of 3 measurements ±Standard error (3 independent experiments).

3.3. Comet assay optimization

Alkaline comet assay is a very sensitive, relatively simple and cheap method for DNA single-strand break detection. The principle of the comet assay involves negatively charged DNA nucleoids that migrate from cathode to anode. Comet slides are then stained with SYBR gold fluorescent dye and viewed under the fluorescent microscope, where represent bright orange spheres (**Figure 3.1. a, b**) if not damaged and tailed comets (**Figure 3.1. c, d**) if strand breaks occurred.

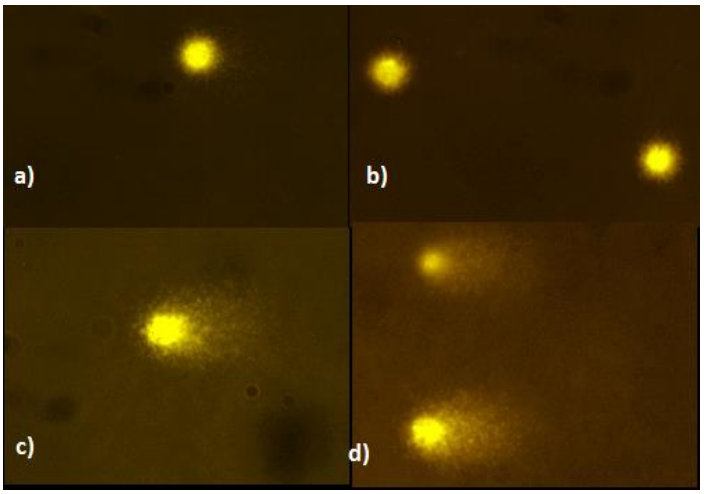


Figure 3.1. Undamaged cells a), b) and cells with DNA damage c), d) under fluorescent microscope after Comet assay. During electrophoresis DNA migrates from anode to cathode, creating comet-like structures, if DNA strand breaks occurred (pictures made using AVT Smartview software).

3.3.1. Stage 1. Preliminary data.

First comet assay stage was performed according to the general protocol. For the positive control 50 µM H_2O_2 (1:2000 in DMEM) was used to verify that the assay is working and DNA damage occurs. After the treatment cells were scraped in PBS and then the rest of assay steps were carried out through all study without modifications.

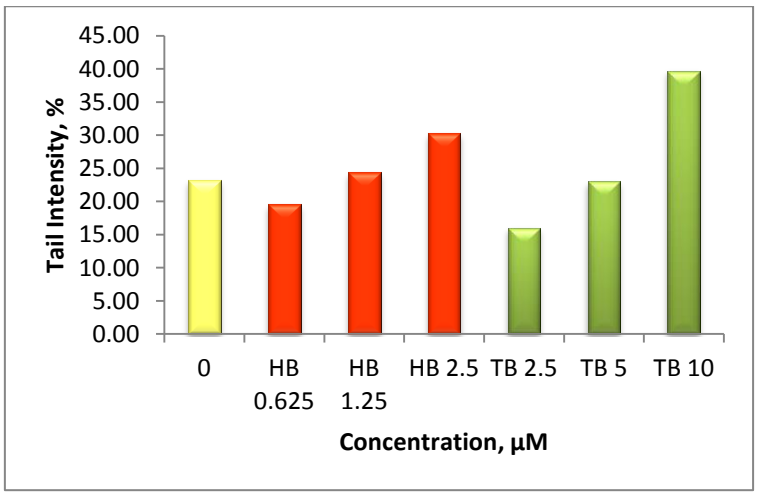


Chart 3.10. BFR caused DNA damage in SH-SY5Y cells. Preliminary data. Comet tail intensity (%), which indicates the percentage of DNA single-strand breaks was measured following 24 hours treatment with HBCD (0-2.5 µM) and TBBP-A (0-10 µM) employing Comet assay. Each data point is a mean value of 100 measurements (cells recorded).

The results of the first comet assay stage (**Chart 3.10**) indicated damaged negative control (tail intensity – 23.2%, while normally is only 2-4%) and very high overall damage in all samples. Moreover, in the positive control as well as in 2 tested concentrations (5 µM for HBCD and 20 µM for TBBP-A, later excluded as in previous assays) DNA damage was undetectable – during the analysis only random DNA fragments were identified and no comets were observed. However, at this stage a concentration-dependent increase was already evident for 2 compounds – HBCD and TBBP-A.

3.3.2. Stage 2. Optimization.

Optimization step was carried out to reduce the high damage in samples detected at the first stage. For better performance the protocol was modified for SH-SY5Y cell type. The results showed that SH-SY5Y neuroblastoma cells are highly sensitive to scraping, but are easy trypsinised without detectable damage. Therefore, further experiments were performed with

trypsin instead of scraping. Two dilutions were carried out to reduce the damage in positive control – 25 mM of H_2O_2 (1:4000 in DMEM) and 12.5 mM of H_2O_2 (1:8000 in DMEM). Still, the cells were too damaged after an hour treatment – comets were too fragmented, and as a result excluded from the analysis.

3.4. Alkaline Comet assay

Optimization step noticeably improved comet assay performance. The negative control value significantly declined (being within norm – 2.5%) and DNA damage within samples decreased as well.

Results of the comet assay indicated same trend with all compounds – DNA damage considerably increased in a concentration-dependent manner (**Chart 3.11-3.13**). Even at lowest concentration the percentage of single-strand breaks doubled with all three compounds. HBCD was the most potent chemical in the assay, increasing almost 6 fold at the highest concentration used (5 µM) in comparison to the control and causing considerable DNA damage in cells.

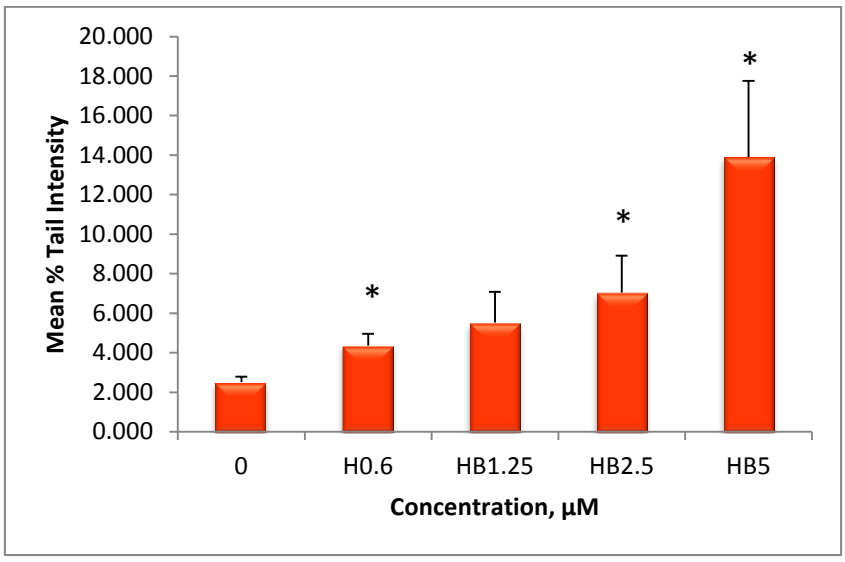


Chart 3.11. DNA damage in HBCD treated SH-SY5Y cells. Comet tail intensity (%) was measured following 24 hours treatment with HBCD (0-5 µM) using comet assay. Each data point represents a

mean value of 300 measurements ±Standard error (3 independent experiments). *Asterisks indicate significantly different values ($p<0.05$) calculated by Student's T-test.

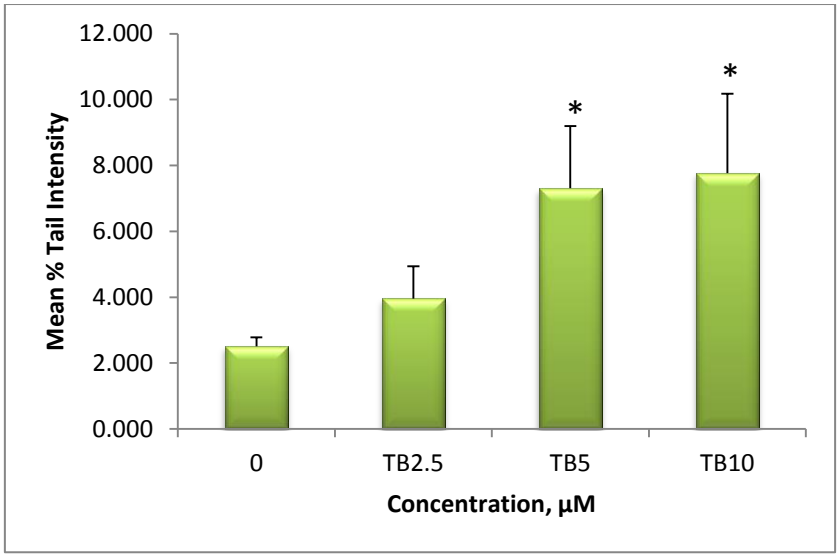


Chart 3.12. DNA damage in TBBP-A treated cells. Comet tail intensity (%) was measured following 24 hours treatment with TBBP-A (0-10 µM) using comet assay. Each data point represents a mean value of 300 measurements ±Standard error (3 independent experiments). *Asterisks indicate significantly different values ($p<0.05$) calculated by Student's T-test.

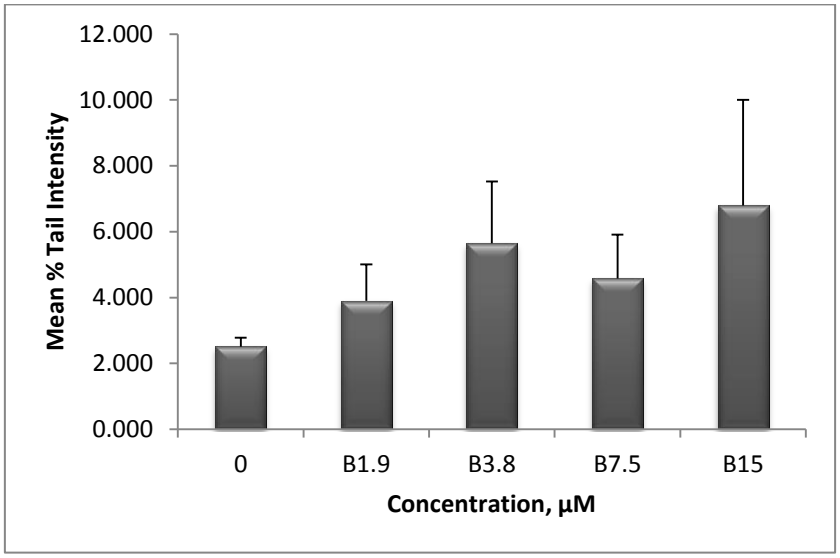


Chart 3.13. DNA damage in BDE-209 treated cells. Comet tail intensity (%) was measured following 24 hours treatment with BDE-209 (0-15 µM) using comet assay. Each data point represents a mean value of 300 measurements ±Standard error (3 independent experiments).

3.5. Apoptosis assay. Preliminary results

Apoptotic or necrotic cell death pathway was determined using vybrant apoptosis assay. The type of the cell death was identified employing FITC annexin V and Propidium iodide labeling. Annexin V is an anticoagulant that binds to phosphatidylserine, which in apoptotic cells is transferred to the outer membrane layer (in live cells it is found on cytoplasmic surface). Annexin contains fluorophore or biotin that labels apoptotic cells. PI dye in contrast labels necrotic cells, as passing through the damaged membrane binds directly to the nucleic acids. Cells stained with the reagents express green fluorescence, if they are apoptotic, red and green fluorescence being necrotic and show almost no fluorescence, if they are alive.

To evaluate the efficiency of the assay in different environment compounds were diluted in 2 medias – serum free and serum present DMEM. Three parameters of the assay were defined and analyzed (**Figure 3.2-3.4**) – percentage of cellular debris (fluorescent objects smaller in size than the cell), PI and FITC staining. The percentage of cellular debris (**Chart 3.14**) increased in cells treated in serum present media with HBCD and TBBP-A compared to the serum free cells, but decreased with BDE-209. The highest percentage of debris was observed in cells exposed to HBCD.

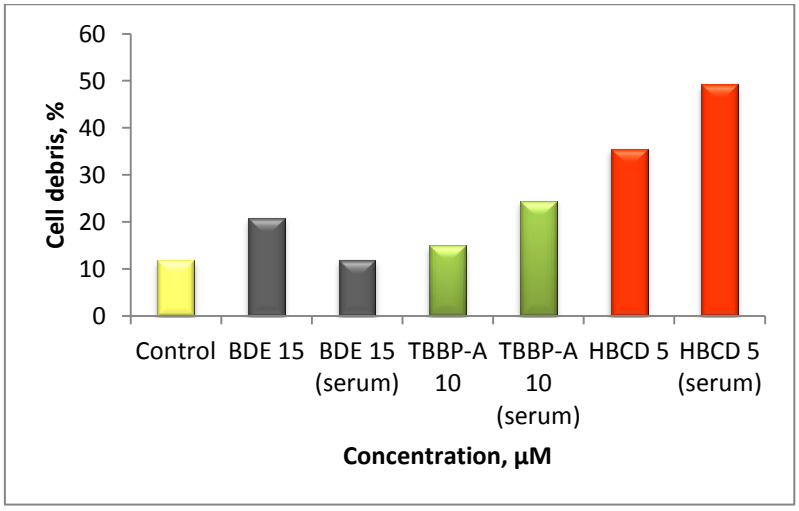


Chart 3.14. Percentage of cell debris in BFR treated SH-SY5Y cells. Following 24 hours treatment with HBCD (5 µM), TBBP-A (10 µM) and BDE-209 (15 µM) in presence and absence of serum

percentage of fragments smaller than cell size was measured by flow cytometry. Necrotic cell fragments and apoptotic bodies due to their size are not present in FITC and PI analysis as are not whole cells, therefore are often quantified as cell debris.

The percentage of necrotic cells was determined using PI labeling (**Chart 3.15**). The highest amount of cells stained with PI was observed with HBCD. The percentage of PI labeled cells was greater with serum in HBCD and TBBP-A samples and smaller in BDE-209. In contrast, the amount FITC stained cells (**Chart 3.16**) considerably increased in serum free media, with TBBP-A (almost twice) and especially with HBCD (rising almost 6 times). BDE-209 again showed the different trend decreasing in serum free media.

Overall, all compounds showed increased FITC annexin V staining in comparison to the control. HBCD proved to be the most potent also demonstrating increased Propidium iodide labelling.

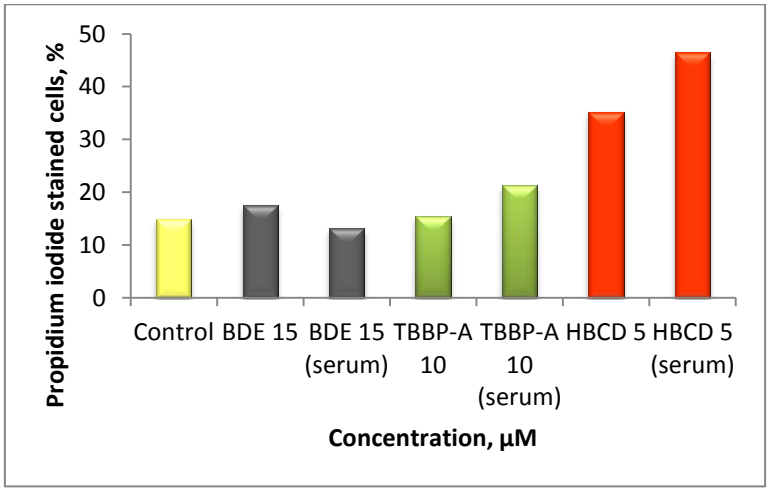


Chart 3.15. Percentage of necrotic cells after BFR treatment. Following 24 hours treatment with HBCD (5 µM), TBBP-A (10 µM) and BDE-209 (15 µM) in presence and absence of serum percentage of PI stained cells was measured by flow cytometry.

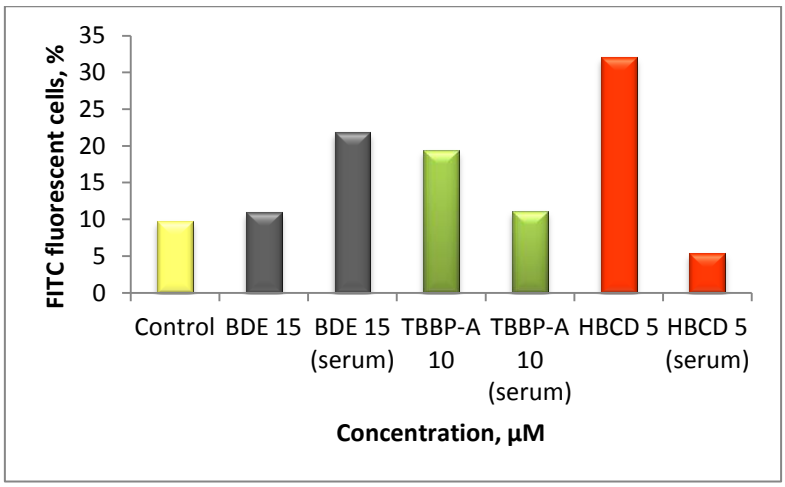


Chart 3.16. Percentage of apoptotic cells after BFR treatment. Following 24 hours treatment with HBCD (5 µM), TBBP-A (10 µM) and BDE-209 (15 µM) in presence and absence of serum percentage of FITC annexin V fluorescent cells was measured by flow cytometry.

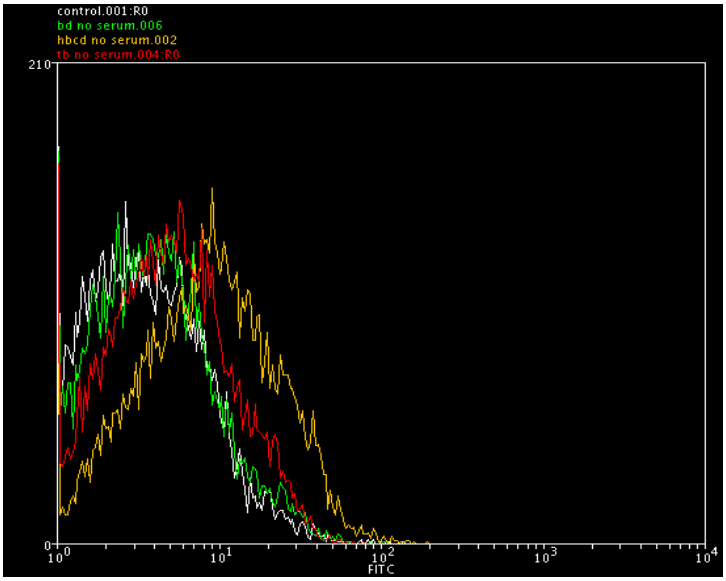


Figure 3.2. FITC fluorescence in SH-SY5Y (serum-free) cells. Lines represent negative control cells (white), cells treated with HBCD (orange), TBBP-A (red) and BDE-209 (green).

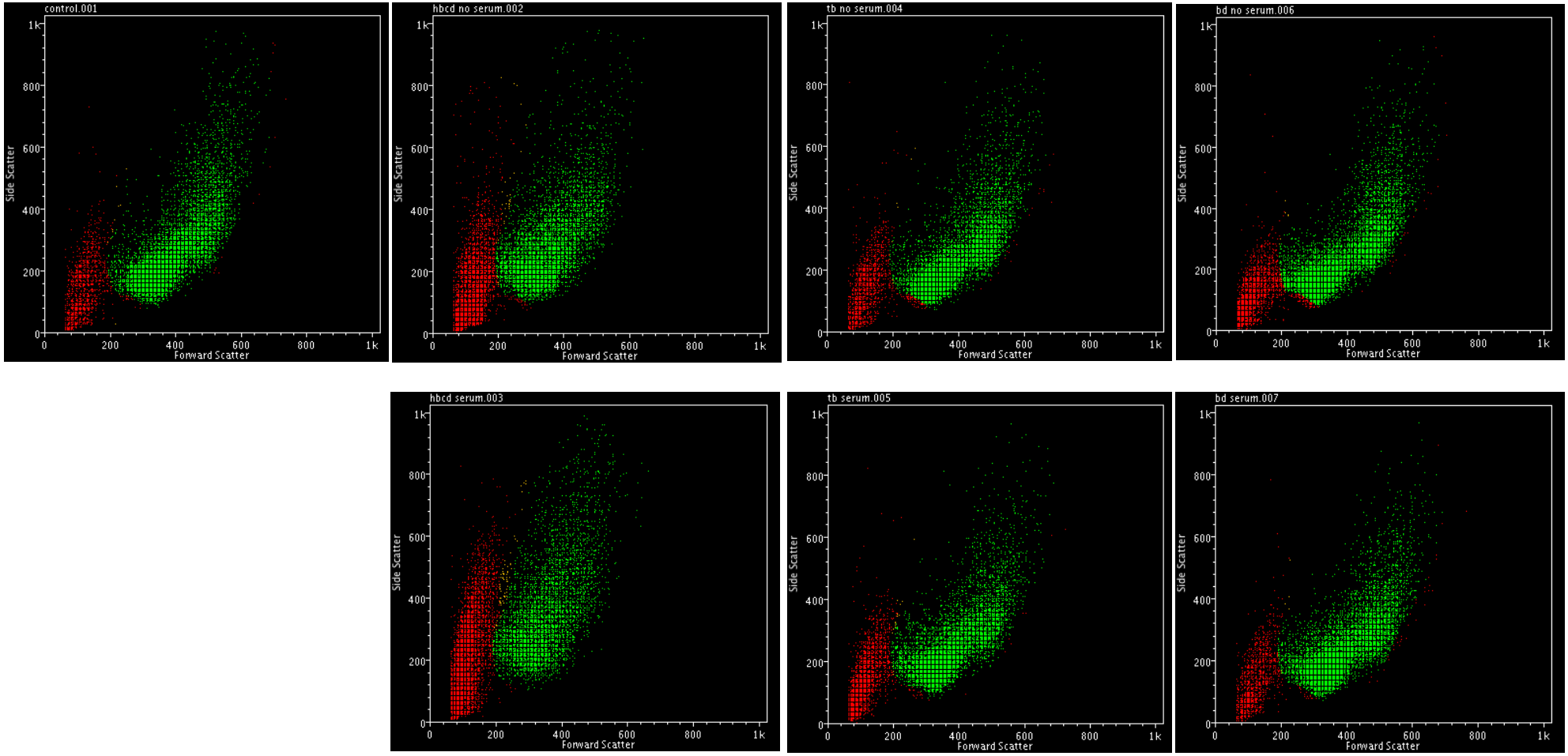


Figure 3.3. Cellular debris in SH-SY5Y cells. Green color represent viable cells, red – cellular debris in negative control cells, cells treated with HBCD, TBBP-A and BDE-209 in serum-free (top) and serum-present (bottom) medias.

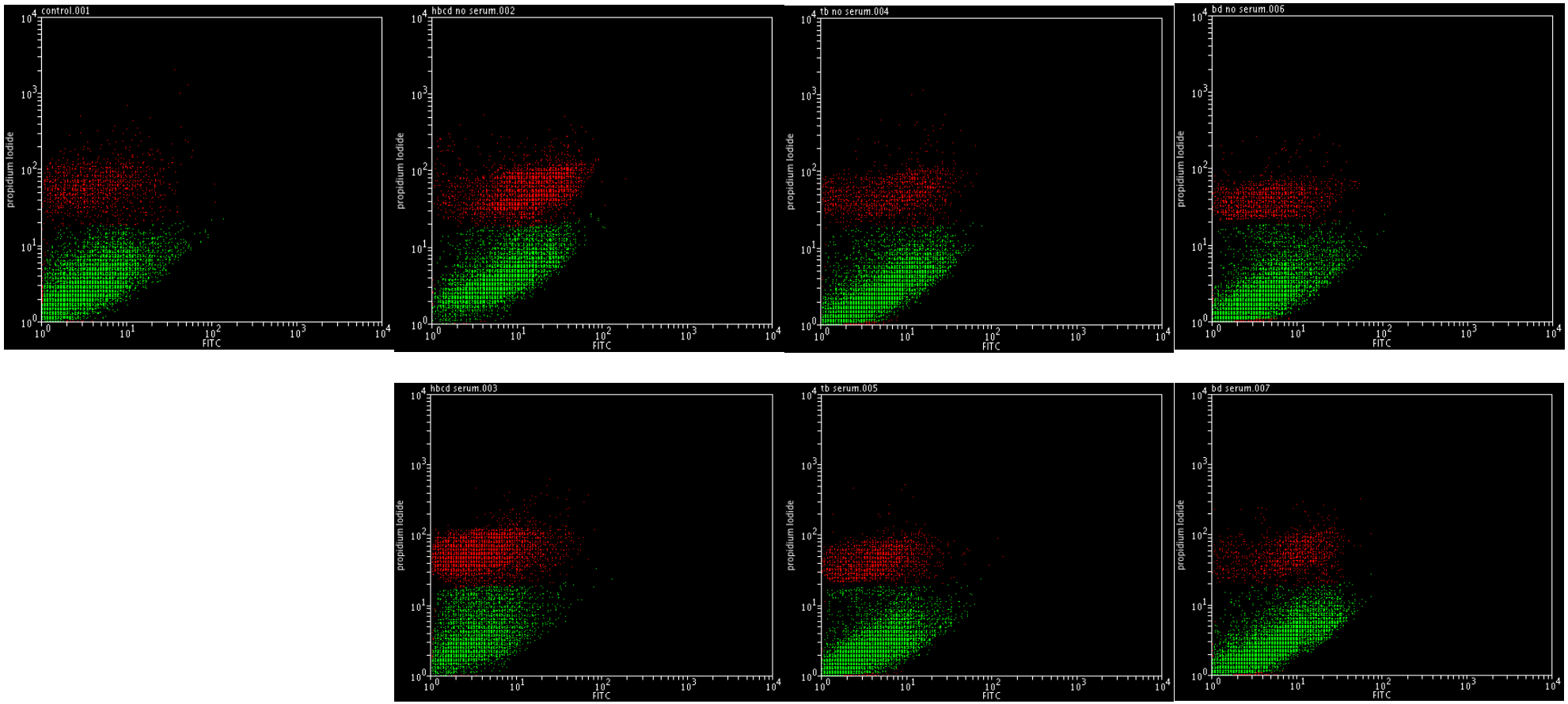


Figure 3.4. Propidium iodide staining in SH-SY5Y cells. Green color represent viable cells, red –PI labelled (necrotic) in negative control, cells treated with HBCD, TBBP-A and BDE-209 in serum-free (top) and serum-present (bottom) medias.

3.1. Autophagy. Preliminary results

Recombinant DNA technology using the transformation process involves the formation of selected DNA fragment clones. The fragment required is bonded to *E.coli* plasmid vector, replicative inside a host cell. This recombinant DNA molecule, containing drug-resistance gene, Enhanced Green Fluorescent Protein (EGFP) gene and LC3 protein (rat microtubule-associated protein 1 light chain 3) inserted into the cell produces millions of copies of the required fragment (Lodish et al., 2000). Following general protocol, *E. coli* Super Competent cells were successfully transformed using pEGFP-C1-LC3 vector. Plasmid DNA was then isolated from three *E. coli* transformed colonies and isolation efficiency was then monitored measuring DNA absorbance (260 nm) for all clones obtained (being 0.217, 0.368 and 0.337 µg/µl for Clone 1, 2 and 3 respectively).

After SH-SY5Y cells were transfected (**Figure 3.5. a, b**) with plasmid DNA, employing TurboFect reagent – cationic polymer, that creates stable and positively charged structures with DNA. Transfection efficiency was examined under fluorescent microscope. Around 60-70% of cells expressed green fluorescent protein (no fluorescence observed in negative control), what makes the results of transfection very effective. Transfection reagent however was slightly toxic for the cells; nevertheless in general cell viability was not considerably affected.

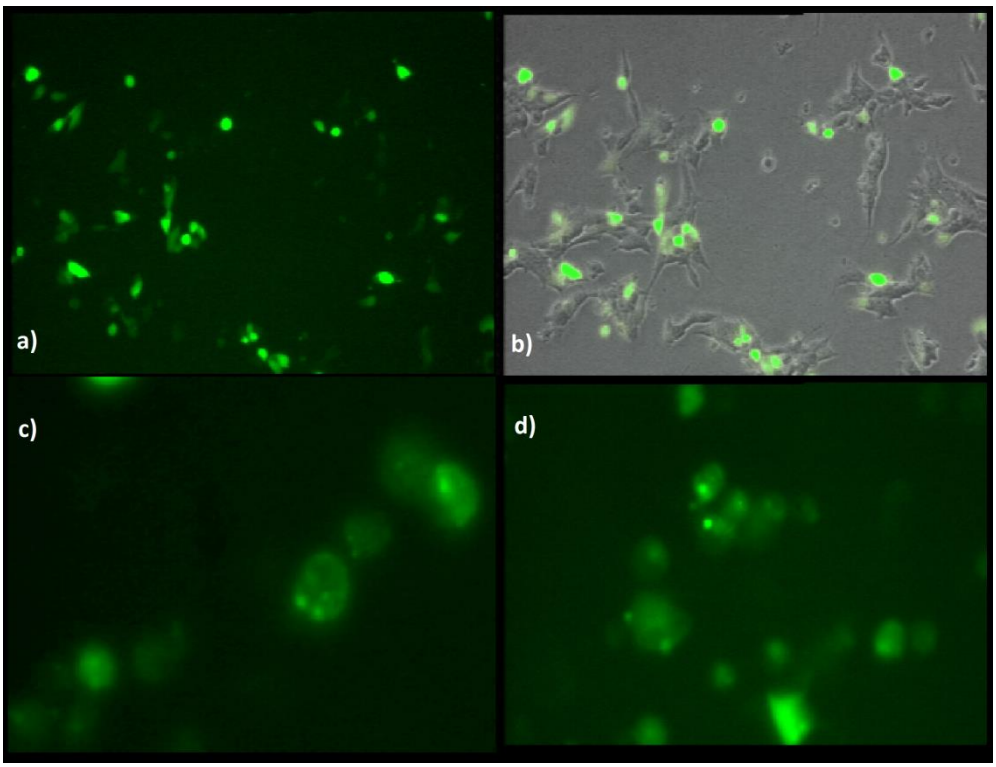


Figure 3.5. GFP expression a), b) and cytoplasmic puncta (autophagy) c), d) in transfected SH-SY5Y cells.

After 48 hours transfected cells, treated with compounds, were fixed and monitored under the fluorescent microscope. Autophagic process was examined in GFP positive, treated cells. The principle of autophagy assay involves GFP-LC3 fusion gene expression that permits to visualize the development of autophagic vesicles in cells. At the stage of autophagosome formation GFP-LC3 binds to the autophagosome membrane and can be observed as a cytoplasmic puncta (**Figure 3.5. c, d**).

In preliminary experiment (using standard media) after 0 and 4 hours treatment no punctate cells were detected in samples. In 24 hours some cells were punctate in samples with HBCD and TBBP-A, however generally no significant changes were observed. Nevertheless, in another autophagy experiment in cells treated with HBCD (diluted in clear serum-free media) only for 2 hours considerable changes were detected. Percentage of cytoplasmic puncta increased from 21 to 50% compared to the control.

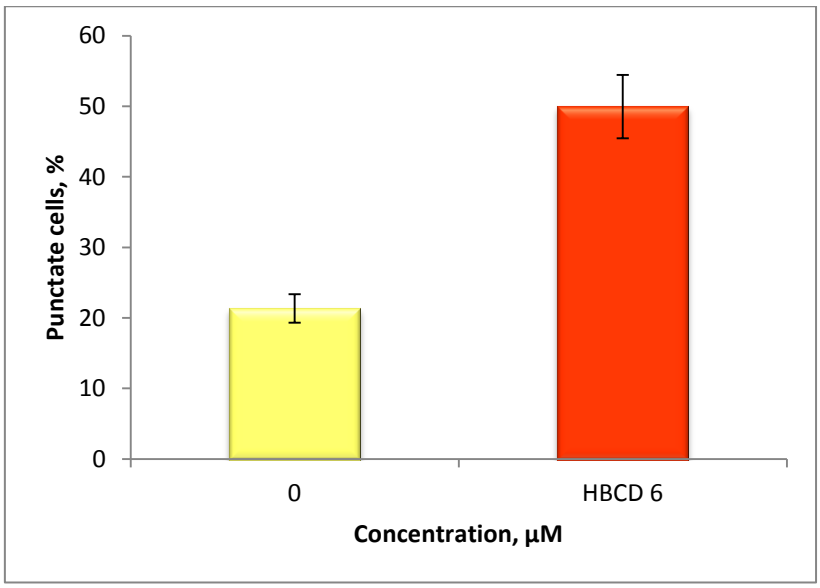


Chart 3.17. Autophagic puncta in SH-SY5Y cells treated with HBCD. Cells were treated with 0 and 6 µM HBCD for 2 hours. Autophagic puncta was calculated in 19 quadrants (103 cells in total) for the control and in 23 (98 cells in total) for HBCD. Data points represent percentage of autophagy in cells ± Standard error.

4. Discussion

During the last decade growing consumption of BFRs raised the concern of their toxic impact on the environment and human health. BFR effects on nervous system and neuronal development have been in focus for the past few years as some studies indicated BFR ability to cross the blood barrier and accumulate within brain tissue (Szabo et al., 2010). The neurotoxic effects observed after the exposure include aberrations in spontaneous behaviour, learning and memory function (Eriksson, 2002; Costa & Giordano, 2011). *In-vitro* studies showed that HBCD, TBBP-A and BDE-209 reduce cell viability, increase ROS formation and elevate [Ca^{2+}] levels in neuronal cells, as well as cause depolarization of the mitochondria and cytochrome c release (Al-Mousa & Michelangeli, 2012). However, the information regarding *in-vitro* neurotoxicity of HBCD, TBBP-A and BDE-209 is still scarce. Therefore, in this report the ability of these three compounds to decrease cell viability, cause oxidative stress and DNA damage was evaluated. Additionally, cell death mechanisms were also studied.

4.1. BFR cytotoxicity

In current report MTT assay results indicated that HBCD, TBBP-A and BDE-209 induce concentration-dependent decrease in cell viability, with HBCD being the most potent of all compounds (LC50 between 2.5 and 5 µM, 15 and 20 µM, 30 and 60 µM for HBCD, TBBP-A and BDE-209 respectively). Recent study confirmed depletion in cell viability in SH-SY5Y cells with all three chemicals after 24 hours exposure and also proved that HBCD (0-30 µM range used) is the most toxic compound. Similar LC50 values were identified being 2.7±0.7 µM for HBCD, 15±4 µM for TBBP-A and 28±7 µM for BDE-209 (Al-Mousa & Michelangeli, 2012). In previous study on TBBP-A (using 0-50 µM range) gradual decrease

in cell viability after 18 hours in mouse TM4 Sertoli cells was detected, with LC50 being 18 ± 6 µM (Ogunbayo et al., 2008). Other reports showed that in 24 hours BDE-209 decreases cell viability in rat hippocampal neurons at the concentrations of 10, 30 and 50 µM (Zhang et al., 2010; Chen et al., 2010).

4.2. Oxidative stress

Following the exposure to various compounds, for instance BFRs, elevated levels of reactive oxygen species are formed, what leads to oxidative stress. Consequently, ROS levels are one of the main oxidative stress parameters (Siddique, 2008). The results of this research showed that concentration-dependent increase in reactive oxygen species occurs only in cells exposed to TBBP-A, what indicates that TBBP-A is able to cause oxidative stress at non cytotoxic concentrations. The most recent study, however, demonstrated that in SH-SY5Y cells at similar concentrations HBCD (1, 3, 5, and 15 µM) as well as TBBP-A (5 µM) cause gradual increase in ROS (after 24 hours). BDE-209 (5 µM) in the same study showed no significant changes in ROS in comparison to the control (Al-Mousa & Michelangeli, 2012). Latest *in-vivo* research in rare minnow (*Gobiocypris rarus*) indicated that 28 and 42 day exposure to waterborne HBCD induces elevation of ROS (in brain) at 0.001, 0.01, 0.1 and 0.5 µM concentrations (Zhang et al., 2008). Studies in rat hippocampal neurons showed that BDE-209 considerably increases ROS levels at 10, 30 and 50 µM concentrations in 24 hours (Zhang et al., 2010; Chen et al., 2010).

Antioxidants and in particular GSH are the main cell defense system that remove ROS from the cell and prevent oxidative stress (Tocher et al., 2002). Therefore, decrease in GSH and elevated ROS formation – two main indicators of oxidative stress are closely connected (Chuang & Chen, 2004). In current study the reduction of GSH in a concentration-dependent manner was observed only with HBCD, but not in cells treated with TBBP-A. That is rather

puzzling as makes 2 major parameters of oxidative stress – increase in ROS and depletion of GSH unrelated in current experiments. This may occur due to lowered assay sensitivity or errors, or the probability that compounds induce toxic reaction through other mechanisms, therefore future experiments are required. *In-vivo* research in rare minnow (*Gobiocypris rarus*) showed that waterborne HBCD (at 0.001, 0.01, 0.1 and 0.5 µM concentrations) after 28 and 42 day exposure induces significant depletion of glutathione (Zhang et al., 2008).

4.3. Genotoxicity

DNA damage is a crucial factor in many human disorders including genetic diseases and cancer. DNA strand breaks can be induced by several factors such as increase in reactive oxygen species, UV radiation or various chemicals (McKenna et al., 2008). Comet assay was created as a one of the most sensitive methods for DNA single-strand break detection in individual cells (Collins, 2008; McKenna et al., 2008; Tice et al., 2008). It is referred to as a rather fast method suitable for all animal cells. However, the assay is not standartized, therefore optimization and analysis is time consuming and can involve several modifications for a particular cell type, also the process consists of many stages (requires a huge effort and preciseness). In addition, the assay is unable to identify small DNA fragments, mitochondrial DNA and apoptotic cells as they are washed during the lysis and electrophoresis (Nossoni, 2008). The problems during the experiment often include too damaged negative control, absence of tails or undetectable comets in positive control, or bubbles in agarose. Nevertheless, after successful optimization stage the assay is able to generate excellent results.

Modified comet assay showed considerable concentration-dependent increase in DNA single-strand breaks with all compounds tested, especially with HBCD. The most recent study in SK-N-MC human neuroblastoma cells indicated that BDE-209 similarly to its lower

brominated congener BDE-47 (tetrabromodiphenyl ether) cause DNA damage at 5, 10, and 20 µM concentration in 4 and 24 hours (Pellacani et al., 2012), however in current study the trend was not strong enough to suggest that, therefore future experiments are essential. Another study on BDE-47 confirmed that it induces DNA damage in SH-SY5Y neuronal cells at 2 µM concentration, without detectable impact on cell viability (He et al., 2009). An *in-vivo* study in rare minnow showed that waterborne HBCD induces DNA damage (at 0.001, 0.01, 0.1 and 0.5 µM concentrations) in fish blood cells following 28 and 42 day exposure (Zhang et al., 2008). However, no studies investigated the ability of HBCD and TBBP-A to cause DNA strand breaks *in-vitro*, and the results obtained with these two compounds are novel.

4.4. Mechanisms of cell death

To understand the molecular mechanisms of BFRs it is crucial to know by which mechanism they cause cell death. Preliminary results of the current study indicated that cell death induced by BFRs occurs mostly via apoptosis with TBBP-A and BDE-209, whereas HBCD is able to cause cell death via both mechanisms (apoptosis and necrosis). Latest study indicated that all three BFRs (HBCD, TBBP-A, BDE-209) cause caspase-dependent apoptosis in SH-SY5Y neuroblastoma cells (Al-Mousa & Michelangeli, 2012). Previous report on TBBP-A in mouse TM4 Sertoli cells confirmed that the compound induces cell death via apoptosis, however specified the possibility that other cell death mechanisms are employed as well (Ogunbayo et al., 2008). Studies in rat hippocampal neurons indicated that BDE-209 cause apoptotic cell death (Zhang et al., 2010; Chen et al., 2010). Another report on human hepatoma cells HepG2 confirmed that BDE-209 induces cell death by apoptotic mechanism (Hu et al., 2007). One more *in-vivo* research in zebrafish embryos also showed that HBCD cause cell death via apoptosis (Deng et al., 2009). To confirm the results obtained in the preliminary experiments of this report additional research is needed.

Autophagy – a complex catabolic process for lysosomal degradation of proteins and other subcellular constituents is usually initiated as a response to nutrient deficiency. It is a key factor in aging, cancer and neurodegeneration prevention, as well as in elimination of intracellular pathogens (Dunn, 1994). Mammalian protein LC3 plays the main role in autophagic assays as associates with autophagosome membrane and in complex with GFP is detectable as fluorescent puncta (Kabeya et al., 2000).

For autophagy monitoring 60-70% transfection efficiency was reached, what is similar to previous studies in human cervical carcinoma HeLa cells, transfected for 12 hours with 70% transfection efficiency obtained (Criollo et al., 2010). So far no studies assessed the ability of brominated flame retardants to induce autophagy, however previous report on other flame retardant group – polychlorinated biphenyls (PCB) indicated that PCBs cause deregulated autophagy in sheep blastocysts (Ptak et al., 2012). The autophagy experiment indicated that HBCD can induce autophagy in SH-SY5Y cells, therefore suggesting that all three mechanisms are employed in HBCD toxicity-caused cell death. However, further research is required to confirm the results obtained with HBCD and to determine if autophagy occurs with TBBP-A and BDE-209 as well.

5. Conclusions

The molecular mechanisms of neurotoxicity by three brominated flame retardants – HBCD, TBBP-A and BDE-209 were explored in this study. The results showed that all three compounds and, especially HBCD, are cytotoxic for SH-SY5Y cells as they decreased cell viability in μM concentrations. Cell death induced by HBCD, TBBP-A and BDE-209 occurs mainly via apoptotic mechanism, whereas HBCD is also able to induce cell death via necrosis and autophagy. However, further research on cell death mechanisms is required to confirm this preliminary data. Following the results of fluorescence spectroscopy increased ROS formations were observed only with TBBP-A and depletion of GSH was detected only with HBCD. From this data it can be concluded that either the assay sensitivity is low, optimization of the assay is needed or another mechanisms are employed in oxidative stress caused by BFRs. Novel comet assay results indicated that all compounds cause DNA damage in human neuroblastoma cells as considerably increase the percentage of single-strand breaks at non-toxic concentrations. The highest DNA damage is caused by HBCD, whereas BDE-209 is least potent due to its low solubility. Therefore, it can be concluded that all three BFRs are not only neurotoxic, but also genotoxic at very low concentrations in SH-SY5Y cell *in-vitro*.

Appendix I. Conferences and publication

This work was presented on British Toxicology Society Annual Congress in Solihull (United Kingdom) in April 2013 under the following title: "Induction of oxidative stress and single strand DNA breaks by brominated flame retardants in SHSY-5Y neuronal cells".

It was also featured (as a poster) in 46th Congress of the European Societies of Toxicology, EUROTOX in September 2013 in Interlaken Switzerland under following title: "An investigation into the toxicity and genotoxicity of brominated flame retardants in SHSY-5Y cells".

Consequently the abstract from EUROTOX congress was also published in Toxicology Letters: Volume 221S, S140, P12-11, August 2013; and can be viewed, following the link: http://dx.doi.org/10.1016/j.toxlet.2013.05.278

References

1. Abdallah M, Harrad S, Covaci A. (2008) Hexabromocyclododecanes and tetrabromobisphenol-A in indoor air and dust in Birmingham, UK: Implications for human exposure. Environ. Sci. Technol., 42: 6855–6861
2. ACC 1. (2001) Data Summary and test plan for hexabromocyclododecane (HBCD). AR201-13459A, American Chemistry Council: Arlington, VA
3. ACC 2. (2001) Data Summary and test plan for Tetrabromobisphenol A, (TBBPA). CAS No. 79-94-7, American Chemistry Council: Arlington, VA
4. Alaee M, Arias P, Sjodin A, Bergman A. (2003) An overview of commercially used brominated flame retardants, their applications, their use patterns in different countries/regions and possible modes of release. Elsevier Science, Environment International, 29: 683– 689
5. Al-Mousa F. and Michelangeli F. (2012) Some Commonly Used Brominated Flame Retardants Cause Ca2+-ATPase Inhibition, Beta-Amyloid Peptide Release and Apoptosis in SH-SY5Y Neuronal Cells. PLoS ONE, 7(4): 1-8
6. AMAP, Arctic Monitoring Assessment Program. (2009) Arctic Pollution: Oslo, Norway
7. Bejarano E. and Cuervo A. (2010) Chaperone-mediated autophagy. Proc. Am. Thorac. Soc., 7: 29–39
8. Birnbaum L. and Staskal D. (2004) Brominated flame retardants: cause for concern? Environ. Health Perspect, 112: 9–17
9. Bradford M. (1976) A dye binding assay for protein. Anal. Biochem., 72: 248-254
10. BSEF. (2000) Bromine Science and Environmental Forum, Brussels, Belgium. Data reported at http://205.232.112.21/bsef/docs/Major_Brominated.doc
11. Chen D, Hale RC. (2010) A global review of polybrominated flame retardant contamination in birds. Environ. Int., 36: 800–11
12. Chen J, Liufu C, Sun W, Sun X, Chen D. (2010) Assessment of the neurotoxic mechanisms of decabrominated diphenyl ether (PBDE-209) in primary cultured neonatal rat hippocampal neurons includes alterations in second messenger signaling and oxidative stress. Toxicol. Lett., 192: 431–439
13. Chen N. and Karantza-Wadsworth V. (2009) Role and regulation of autophagy in cancer. Biochim. Biophys. Acta, 1793: 1516–1523

14. Chuang J, Chen T. (2004) Effect of melatonin on temporal changes of reactive oxygen species and glutathione after MPP(+) treatment in human astrocytoma U373MG cells. J. Pineal. Res., 36(2): 117-25
15. Clarke P. (2002) Apoptosis: from morphological types of cell death to interacting pathways. Trends Pharmacol. Sci., 23: 308–309
16. Collins A. (2008) The comet assay for DNA damage and repair: principles, applications, and limitations. Mol. Biotechnol., 26: 249-261
17. Costa Lucio G. and Giordano G. (2007) Developmental neurotoxicity of polybrominated diphenyl ether (PBDE) flame retardants**.** NeuroToxicology, 28: 1047–1067
18. Costa Lucio G. and Giordano G. (2011) Is decabromodiphenyl ether (BDE-209) a developmental neurotoxicant? NeuroToxicology, 32: 9–24
19. Covaci A, Gerecke A, Law R, Voorspoels S, Kohler M, Heeb N, Collin R. Allchin H, and de Boer J. (2006) Hexabromocyclododecanes (HBCDs) in the Environment and Humans: A Review. American Chemical Society, Environmental science & technology, 40(12): 3679–3688
20. Covaci A, Voorspoelsb S, Abou-Elwafa Abdallah M, Geensa T, Harradc S, Law R. (2009) Analytical and environmental aspects of the flame retardant tetrabromobisphenol-A and its derivatives. J. Chromatogr., 1216: 346–363
21. Criollo A, Senovilla L, Authier H, Maiuri M, Morselli E, Vitale I et al. (2010) The IKK complex contributes to the induction of autophagy. The EMBO Journal, 29: 619–631
22. Darnerud P. (2003) Toxic effects of brominated flame retardants in man and in wildlife. Environ. Int., 29: 841-853
23. De Wit C. (2002) An overview of brominated flame retardants in the environment. Chemosphere, 46: 583-624
24. De Wit C, Alaee M, Muir D. (2004) Brominated flame retardants in the Arcticsan overview of spatial and temporal trends. Organohalogen. Compd., 66: 3811-3816
25. Deng J, Yu L, Liu C, Yu K, Shi X, Yeung L, Lam P, Wu R, Zhou B. (2009) Hexabromocyclododecane-induced developmental toxicity and apoptosis in zebrafish embryos. Aquatic Toxicology, 93: 29–36
26. Dingemans M, Heusinkveld H, de Groot A, Bergman A, van den Berg M, Westerink R. (2009) Hexabromocyclododecane inhibits depolarization-induced increase in intracellular calcium levels and neurotransmitter release in PC12 cells. Toxicol. Sci., 107: 490–497

27. Dunn W. (1994) Autophagy and related mechanism of lysosome-mediated protein degradation. Trends Cell Biol., 4: 139-143
28. ECRA. (2008) European Commission Risk Assessment Hexabromocyclododecane. EINECS No: 247-148-4, Scientific Committee on Health and Environmental Risks: Brussels, Belgium
29. Eguchi Y, Srinivasan A, Tomaselli K, Shimizu S, Tsujimoto Y. (1999) ATP-dependent steps in apoptotic signal transduction. Cancer Res., 59(9): 2174–2181
30. EHC-172, Environmental Health Criteria 172. (1995) Tetrabromobisphenol A and Derivatives. International Programme on Chemical Safety, World Health Organization: Geneva, Switzerland
31. EHC-192. (1997) Flame-retardants: a general introduction. International Program on Chemical Safety, World Health Organization: Geneva, Switzerland
32. Ellis R, Yuan J, Horvitz H. (1991) Mechanisms and functions of cell death. Annu. Rev. Cell Biol., 7: 663–698
33. ERA. (2008) European Risk Assessment Report on 2,2',6,6'-tetrabromo-4,4'-isopropylidenediphenol(tetrabromobisphenol-A or TBBP-A). Part I, Environment, European Commission, Joint Research Centre, European Chemicals Bureau, in draft
34. Eriksson P, Viberg H, Fischer C, Wallin M, Fredriksson A. (2002) A comparison on developmental neurotoxic effects of hexabromocyclododecane, 2,2′,4,4′,5,5′-hexabromodiphenylether (PBDE 153) and 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB 153). Organohalogen. Compd., 57: 389-392
35. Festjens N, Vanden Berghe T, Vandenabeele P. (2006) Necrosis, a well-orchestrated form of cell demise: Signalling cascades, important mediators and concomitant immune response. Biochimica et Biophysica Acta, 1757: 1371–1387
36. Fiers W, Beyaert R, Declercq W, Vandenabeele P. (1999) More than one way to die: apoptosis, necrosis and reactive oxygen damage. Oncogene, 18: 7719–7730
37. Fink U, Hajduk F, Wei Y, Mori H. (2008) Flame retardants. SRI Consulting, Specialty Chemicals
38. Fonnum F. and Mariussen E. (2009) Mechanisms involved in the neurotoxic effects of environmental toxicants such as polychlorinated biphenyls and Brominated Flame Retardants. International Society for Neurochemistry, J. Neurochem., 111: 1327–1347
39. Germer S, Piersma A, van der Ven L, Kamyschnikow A, Fery Y, Schmitz H, Schrenk D. (2006) Subacute effects of the brominated flame retardants hexabromocyclododecane and tetrabromobisphenol-A on hepatic cytochrome P450 levels in rats. Toxicology, 218: 229-236

40. Gill U, Chu I, Ryan J, Feelry M. (2004) Polybrominated diphenyl ethers: human tissue levels and toxicology. Rev. Environ. Contamin. Toxicol., 183: 55–97
41. Girard P. and Boiteux S. (1997) Repair of oxidized DNA bases in the yeast Saccharomyces cerevisiae. Biochimie, 79: 559–566
42. Goodman J. (2009) Neurodevelopmental effects of decabromodiphenyl ether (BDE-209) and implications for the reference dose. Regul. Toxicol. Pharmacol., 54: 91–104
43. Goosey E, Harrad S, Desborough J, Abdallah M, Roosens L, Covaci A. (2010) Dust from UK primary school classrooms and daycare centers: The significance of dust as a pathway of exposure of young UK children to brominated flame retardants and polychlorinated biphenyls. Environ. Sci. Technol., 44: 4198–4202
44. Hancock J, Desikan R, Neill S. (2001) Role of Reactive Oxygen Species in Cell Signaling Pathways. Biochemical and Biomedical Aspects of Oxidative Modification, 29(2): 345-350
45. Hardy M. (2000) Distribution of decabromodiphenyl oxide in the environment. Organohalogen. Compd., 47: 237–240
46. Hardy M, Banasik M, Stedeford T. (2009) Toxicology and human health assessment of decabromodiphenyl ether. Crit. Rev. Toxicol., 39(S3): 1–44
47. Harrad S, Abdallah M, Covaci A. (2009) Causes of variability in concentrations and diasteromer patterns of hexabromocyclododecanes in indoor dust. Environ. Int., 35: 573–579
48. Harrad S, de Wit C, Abdallah M, Bergh C, Bjorklund J, Covaci A, Darnerud P, de Boer J, Diamond M, Huber S, Leonards P, Mandalakis M, Oestman C, Haug L, Thomsen C, Webster T. (2010) Indoor contamination with hexabromocyclododecanes, polybrominated diphenyl ethers, and perfluoroalkyl compounds: An important exposure pathway for people? Environ. Sci. Technol., 44 (9): 3221–3231
49. He P, Wang A, Xiam T, Gao P, Niu Q, et al. (2009) Mechanism of the neurotoxic effect of PBDE-47 and interaction of PBDE-47 and PCB153 in enhancing toxicity in SH-SY5Y cells. Neurotoxicology, 30: 10–15
50. Held P. (2010) An Introduction to Reactive Oxygen Species: Measurement of ROS in Cells. BioTek Instruments, Inc.: Highland Park, Winooski, Vermont, USA
51. Helleday T, Tuominen K, Bergman A, Jenssen D. (1999) Brominated flame retardants induce intragenic recombination in mammalian cells. Mutat. Res., 439: 137-147
52. Herzke D, Berger U, Kallenborn R, Nygård T, Vetter W. (2005) Brominated flame retardants and other organobromines in Norwegian predatory bird eggs. Chemosphere, 61: 441-449

53. Higuchi Y. (2003) Chromosomal DNA fragmentation in apoptosis and necrosis induced by oxidative stress. Biochem. Pharmacol., 66: 1527–1535
54. Hites R, Foran J, Schwager S, Knuth B, Hamilton M, Carpenter D. (2004) Global assessment of polybrominated diphenyl ethers in farmed and wild salmon. Environ. Sci. Technol., 38: 4945–4949
55. Hou Y, Janczuk A, Wang P. (1999) Current trends in the development of nitric oxide donors. Curr. Pharm. Des., 5(6): 417–471
56. Hu X, Xu Y, Hu Dc. Hui Y, Yang F. (2007) Apoptosis induction on human hepatoma cells HepG2 of decabrominated diphenyl ether (PBDE-209). Toxicol. Lett., 171: 19–28
57. Hughes P, McLellan H, Lowes D, Kan S, Bilmen J, Tovey S, Godfrey R, Michell R, Kirk C, Michelangeli F. (2000) Estrogenic alkylphenols induce cell death by inhibiting testis endoplasmic reticulum Ca2+ pumps. Biochem. Biophys. Res. Commun., 277: 568–574
58. Hunziker R, Gonsior S, MacGregor J, Desjardins D, Ariano J, Friederich U. (2004) Fate and effect of hexabromocyclododecane in the environment. Organohalogen Compd., 66: 2300-2305
59. Hutzinger O, Thoma H. (1987) Polybrominated dibezo-p-dioxins and dibenzofurans: the flame retardant issue. Chemosphere 1987, 16: 1877– 1880
60. Hwang E and Kim G. (2007) Biomarkers for oxidative stress status of DNA, lipids, and proteins in vitro and in vivo cancer research. Toxicology, 229: 1–10
61. IPCS, International Program on Chemical Safety. (1994) Environmental Health Criteria no. 162. Brominated diphenyl ethers. WHO: Geneva, Switzerland
62. IPCS, International Program on Chemical Safety. (1995) Environmental Health Criteria no. 172. Tetrabromobisphenol A and derivates. WHO: Geneva, Switzerland
63. Jakobsson K, Thuresson L, Rylander A, Sjodin L, Bergman H. (2002) Exposure to polybrominated diphenyl ethers and tetrabromobisphenol A among computer technicians. Chemosphere, 46: 709–716
64. Janak K, Covaci A, Voorspoels S, Becher G. (2005) Hexabromocyclododecane in marine species from the western Scheldt estuary: Diastereoisomer- and enantiomer-specific accumulation. Environ. Sci. Technol., 39: 1987–1994
65. Jefferies H, Coster J, Khalil A, Bot J, Mccauley R, Hall J. (2003) Glutathione. Anz. J. Surg., 73: 517–522

66. Jenssen B, Sormo E, Salmer M, Bæk K, Skaare J. (2004) Brominated flame retardants (BFRs) in the Artic marine food chain. Proceedings of the Third International Workshop on Brominated Flame Retardants: Toronto, Canada, 6-9 June, 207-208
67. Johnson-Restrepo B. and Kannan K. (2009) An assessment of sources and pathways of human exposure to polybrominated diphenyl ethers in the United States. Chemosphere, 76: 542–548.
68. Jones D. (2000) Redox Potential of GSH/GSSG Couple: Assay and Biological Significance. Methods of Enzymology, 348: 93-112
69. Kabeya Y, Mizushima N, Ueno T, Yamamoto A, Kirisako T, Noda T, Kominami E, Ohsumi Y, Yoshimori T. (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. EMBO Journal, 19(21): 5720-5728
70. Kerr J, Wyllie A, Currie A. (1972) Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br. J. Cancer, 26(4): 239–257
71. Kitamura S, Kato T, Iida M, Jinno N, Suzuki T, Ohta S, Fujimoto N, Hanada H, Kashiwagi K, Kashiwagi A. (2005) Anti-thyroid hormonal activity of tetrabromobisphenolA, a flame retardant, and related compounds: Affinity to the mammalian thyroid hormone receptor, and effect on tadpole metamorphosis. Life Sciences, 76: 1589-1601
72. Knutsen H, Kvalem H, Thomsen C, Froshaug M, Haugen M, Becher G, Alexander J, Meltzer H. (2008) Dietary exposure to brominated flame retardants correlates with male blood levels in a selected group of Norwegians with a wide range of seafood consumption. Mol. Nutr. Food Res., 52: 217–227
73. Kong W, Kuester R, Gallegos A, Sipes G. (2011) Induction of DNA damage in human urothelial cells by the brominated flame retardant 2,2-bis(bromomethyl)-1,3-propanediol: Role of oxidative stress. Toxicology, 290: 271–277
74. Kuiper R, van den Brandhof E, Leonards P, van der Ven L, Wester P, Vos J. (2007) Toxicity of tetrabromobisphenol A (TBBPA) in zebrafish (Danio rerio) in a partial life-cycle test. Arch. Toxicol., 81: 1
75. Law K, Halldorson T, Danell R, Stern G, Gerwutz S, Alaee M, Marvin C, Whittle D, Tomy G. (2006) Bioaccumulation and trophic transfer of some brominated flame retardants in a Lake Winnipeg (Canada) food web. Environ. Toxicol. Chem., 25: 2177–2186

76. Law K, Halldorson T, Danell R, Stern G, Gewurtz S, Alaee M, Marvin C, Tomy G. (2007) Bioaccumulation and trophic transfer of some brominated flame retardants in a Lake Winnipeg (Canada) food-web. Environ. Toxicol. Chem., 26: 190
77. Law R, Alaee M, Allchin C, Boon J, Lebeuf M, Lepom P, Stern G. (2003) Levels and trends of polybrominated diphenylethers and other brominated flame retardants in wildlife. Environ. Int., 29: 757-770
78. Lee A, Michelangeli F, East J. (1989) Tests for the importance of fluidity for the function of membrane proteins. Biochem. Soc. Trans., 17(6): 962–964
79. Lee J, Giordano S, Zhang J. (2012) Autophagy, mitochondria and oxidative stress: cross-talk and redox signaling. Biochem. J., 441: 523–540
80. Lodish H, Berk A, Zipursky S, et al. (2000) Molecular Cell Biology. 4th edition. New York: W. H. Freeman; Section 7.1, DNA Cloning with Plasmid Vectors. Available from: http://www.ncbi.nlm.nih.gov/books/NBK21498/
81. Lushchak V. and Semchyshyn H. (2012) Oxidative Stress – Molecular Mechanisms and Biological Effects. InTech. Janeza Trdine 9, 51000: Rijeka, Croatia
82. Mariussen E. and Fonnum F. (2002) The effect of pentabromodiphenyl ether, hexabromocyclododecane and tetrabromobisphenol A on dopamine uptake into rat brain synaptosomes. Organohalog. Compd., 57: 395–399
83. Marvin C, Tomy G, Armitage J, Arnot J, McCarty L, Covaci A, Palace V. (2011) Hexabromocyclododecane: Current Understanding of Chemistry, Environmental Fate and Toxicology and Implications for Global Management. Environ. Sci. Technol., 45: 8613–8623
84. McDonald T. (2005) Polybrominated diphenylether levels among United States residents: daily intake and risk of harm to the developing brain and reproductive organs. Integr. Environ. Assess. Manag., 1: 343–354
85. McKenna D, McKeown S, McKelvey-Martin V. (2008) Potential use of the comet assay in the clinical management of cancer. Mutagenesis, 23: 183-190
86. Meerts I, van Zanden J, Luijks E, van Leeuwen-Bol I, Marssh G, Jakobsson E, et al. (2000) Potent competitive interactions of some brominated flame retardants and related compounds with human transthyretin in vitro. Toxicol. Sci., 56: 95 –104
87. Mijaljica D, Prescott M, Devenish R. (2011) Microautophagy in mammalian cells: revisiting a 40-year-old conundrum. Autophagy, 7: 673–682
88. Morck A, Hakk H, Orn U, Klasson Wehler E. (2003) Decabromodiphenyl ether in the rat: absorption, distribution, metabolism, and excretion. Drug. Metab. Dispos., 31: 900–907

89. Morris S, Allchin C, Zegers B, Haftka J, Boon J, Belpaire C, Leonards P, Van Leeuwen S, de Boer J. (2004) Distribution and fate of HBCD and TBBPA brominated flame retardants in North Sea estuaries and aquatic food webs. Environ. Sci. Technol., 38: 5497
90. Mosmann T. (1983) Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. J. Immunol. Methods, 65: 55-63
91. Murphy M. (2009) How mitochondria produce reactive oxygen species. Biochem. J., 417: 1–13
92. Nossoni F. (2008) Single-Cell Gel Electrophoresis (Comet Assay): Methodology, Potential Applications, and Limitations in Cancer Research. MMG 445, Basic Biotechnology eJournal, 4: 30 – 35.
93. Ogunbayo O, Jensen K, Michelangeli F. (2007) The interaction of the brominated flame retardant: Tetrabromobisphenol A with phospholipid membranes. Biochimica et Biophysica Acta, 1768: 1559–1566
94. Ogunbayo O, Lai P, Connolly T, Michelangeli F. (2008) Tetrabromobisphenol A (TBBPA), induces cell death in TM4 Sertoli cells by modulating Ca2+ transport proteins and causing dysregulation of Ca2+ homeostasis. Toxicology in Vitro, 22: 943–952
95. Oppenheim R, Flavell R, Vinsant S, Prevette D, Kuan C, Rakic P. (2001) Programmed cell death of developing mammalian neurons after genetic deletion of caspases. J. Neurosc., 21(13): 4752–4760
96. Pellacani C, Buschini A, Galati S, Mussi F, Franzoni S, Costa L. (2012) Evaluation of DNA Damage Induced by 2 Polybrominated Diphenyl Ether Flame Retardants (BDE-47 and BDE-209) in SK-N-MC Cells. International Journal of Toxicology, 31(4): 372-379
97. Pettigrew A. (1994) Halogenated flame retardants. 4th ed. Kirk–Othmer encyclopedia of chemical technology, 10: 954– 976: New York, Wiley
98. Ptak G, Zacchini F, Czernik M, Fidanza A, Palmieri C, Della Salda L, Scapolo P, Loi P. (2012) A short exposure to polychlorinated biphenyls deregulates cellular autophagy in mammalian blastocyst in vitro. Hum. Reprod., 27(4): 1034-42
99. Pullen S, Boecker R, Tiegs G. (2003) The flame retardants TetrabromobisphenolA and tetrabromobisphenol A-bisallylether suppress the induction of interleukin- 2 receptor alpha chain (CD25) in murine splenocytes. Toxicology, 184: 11–22

100. Reistad T, Mariussen E, Fonnum F. (2002) The effect of brominated flame retardants on cell death and free radical formation in cerebellar granule cells. Organohalog. Compd., 57: 391–394

101. Reistad T, Mariussen E, Ring A, Fonnum F. (2007) In vitro toxicity of tetrabromobisphenol A on cerebellar granule cells: cell death, free radical formation, calcium influx and extracellular glutamate. Toxicol. Sci., 96: 268–278

102. Ronisz D, Finne E, Karlsson H, Forlin L. (2004) Effects of the brominated flame retardants hexabromocyclododecane (HBCDD) and tetrabromobisphenol-A (TBBP-A) on hepaticenzymes and other biomarkers in juvenile rainbow trout and feral eelpout. Aquat. Toxicol., 69: 229-245

103. Roosens L, Abdallah M, Harrad S, Neels H, Covaci A. (2009) Exposure to hexabromocyclododecanes (HBCDs) via dust ingestion, but not diet, correlates with concentrations in human serum: Preliminary results. Environ. Health Perspect., 117: 1707–1712

104. Sahu R, Kaushik S, Clement C, Cannizzo E, Scharf B, Follenzi A, Potolicchio I, Nieves E, Cuervo A, Santambrogio L. (2011) Microautophagy of cytosolic proteins by late endosomes. Dev. Cell, 20: 131–139

105. Sandholm A, Emanuelsson B, Klasson Wehler E. (2003) Bioavailability and half-life of decabromodiphenyl ether (BDE-209) in rat. Xenobiotica, 33: 1149–1158

106. Schecter A, Haffner D, Colacino J, Patel K, Opel M, Birnbaum L. (2010) Polybrominated diphenyl ethers (PBDEs) and hexabromocyclodecane (HBCD) in composite U.S. food samples. Environ. Health Perspect., 118: 357–362

107. Schneider L. and Zhang J. (2010) Lysosomal function in macromolecular homeostasis and bioenergetics in Parkinson's disease. Mol. Neurodegener., 5: 14

108. Sellstrom U, Kierkegaard A, de Wit C, Jansson B. (1998) Polybrominated diphenyl ethers and hexabromocyclododecane in sediment and fish from a Swedish river. Environ. Toxicol. Chem., 17: 1065-1072

109. Shi Z, Wu Y, Li J, Zhao Y, Feng J. (2009) Dietary exposure assessment of Chinese adults and nursing infants to tetrabromobisphenol-A and hexabromocyclododecanes: occurrence measurements in foods and human milk. Environ. Sci. Technol., 43: 4314–4319

110. Siddique H. (2008) Adverse effect of tannery waste leachates in transgenic *Drosophila melanogaster*: role of ROS in modulation of Hsp70, oxidative stress and apoptosis. Appl. Toxicol., 28(6): 734–748

111. Singh N, McCoy M, Tice R, Schneider E. (1988) A simple technique for quantitation of low levels of DNA damage in individual cells. Exp. Cell Res., 175: 184-191

112. Sjödin A, Patterson D, Bergman Å. (2003) A review on human exposure to brominated flame retardants (BFRs)—particularly polybrominated diphenyl ethers (PBDEs). Environ. Int., 29: 829–839

113. Szabo D, Diliberto J, Hakk H, Huwe J, Birnbaum L. (2010) Toxicokinetics of the flame retardant hexabromocyclododecane gamma: effect of dose, timing, route, repeated exposure, and metabolism. Toxicol. Sci., 117: 282–293

114. Thomsen C, Molander P, Daae H, Janak K, Froshaug M, Liane V, Thorud S, Becher G, Dybing E. (2007) Occupational exposure to hexabromocyclododecane at an industrial plant. Environ. Sci. Technol., 41: 5210–5216

115. Tice R, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H, Miyamae Y, Rojas E, Ryu J, Sasaki Y. (2008) Single cell gel/ comet assay: guidelines for in vitro and in vivo genetic toxicology testing. Environ. Mol. Mutagen., 35: 206-221

116. Tocher D, Mourente G, Van Der Eecken A, Evjemo J, Diaz E, Bell J, Geurden I, Olsen Y. (2002) Effects of dietary vitamin E on antioxidant defense mechanisms of juvenile turbot (*ScoPhthalmus maximus* L.), halibut (*HiPPoglossus hiPPoglossus* L.) and sea bream (*SParus aurata* L.). Aquacult. Nutr., 8: 195–203

117. Tomy G, Pleskach K, Oswald T, Halldorson T, Helm P, Marvin C, MacInnis G. (2008) Enantioselective bioaccumulation of hexabromocyclododecane and congener-specific accumulation of brominated diphenyl ethers in an Eastern Canadian Arctic marine food web. Environ. Sci. Technol., 42: 3634–3639

118. USEPA, United States Environmental Protection Agency. (2010) An exposure assessment of polybrominated diphenyl ethers. USEPA: Washington, DC, p.378

119. Van Loon B, Markkanen E, Hubscher U. (2010) Oxygen as a friend and enemy: how to combat the mutational potential of 8-oxo-guanine. DNA Repair (Amst), 9: 604–616

120. Verreault J, Gabrielsen G, Chu S, Muir D, Andersen M, Hamaed A, Letcher R. (2005) Flame retardants and methoxylatedandhydroxylated polybrominated diphenyl ethers in two Norwegian Arctic top predators: glaucous gulls and polar bears. Environ. Sci. Technol., 39: 6021-6028

121. Vorkamp K, Thomsen M, Falk K, Leslie H, Moller S, Sørensen P. (2005) Temporal development of brominated flame retardants in peregrine Falcon (Falco

peregrinus) eggs from South Greenland (1986-2003). Environ. Sci. Technol., 39: 8199-8206

122. Weiss B. (2000) Vulnerability of children and the developing brain to neurotoxic hazards. Environ. Health Perspect., 108: 375–381

123. Wu J, Guan Y, Zhang Y, Luo X, Zhi H, Chen S, Mai B. (2010) Trophodynamics of hexabromocyclododecanes and several other non-PBDE brominated flame retardants in a freshwater food web. Environ. Sci. Technol., 44: 5490–5495

124. Zeiss C. (2003) The Apoptosis-Necrosis Continuum: Insights from Genetically Altered Mice. Vet. Pathol., 40: 481

125. Zhang C, Liu F, Liu X, Chen D. (2010) Protective effect of N-acetylcysteine against BDE-209- induced neurotoxicity in primary cultured neonatal rat hippocampal neurons in vitro. Int. J. Dev. Neurosci., 28: 521–528

126. Zhang X, Yang F, Zhang X, Xu Y, Liao T, Song S, Wang J. (2008) Induction of hepatic enzymes and oxidative stress in Chinese rare minnow (*Gobiocypris rarus*) exposed to waterborne hexabromocyclododecane (HBCDD). Aquat. Toxicol., 86: 4–11

The University of Birmingham

School of Biosciences

WHOLE-GENOME SEQUENCING, FINISHING AND ANALYSIS OF THREE BACTERIAL HUMAN PATHOGENS

A research project report submitted by

Jelena Sostare

as a part of the requirement for the

degree of MRes in Molecular and Cellular Biology

Project supervisor: Dr Nick Loman

Birmingham, 2013

Abstract

Nowadays, multidrug resistant bacteria are one of the greatest human concerns as they commonly cause illness and death. A new widely-used technique – high-throughput (HTS) whole-genome sequencing, allows rapid genome finishing thus providing all of the resistance information (and more), and helps to find the appropriate cure or even prevent bacterial outbreaks rapidly and in just one step. Therefore, this study attempted to finish the genome of multidrug-resistant *Elizabethkingia meningoseptica* 501 (induce meningitis in infants) and start finishing the genome of a new *Pseudomonas aeruginosa* ST395 strain (caused an outbreak in Birmingham hospital in burn victims), employing whole-genome Nextera XT and Mate Pair sequencing (Illumina). Optimization of Nextera XT for organisms with different GC content was also carried out in the study using 3 species – *Escherichia coli* (Medium GC), *Pseudomonas aeruginosa* (high GC) and *Elizabethkingia meningoseptica* (low GC). Nextera XT results indicated a concentration-dependant increase in *E. coli* samples and allowed to obtain contiguous sequences for each genome. Furthermore, Mate Pair sequencing enabled joining contigs into scaffolds. PCR reactions and Sanger sequencing allowed closing sequencing gaps and start finishing the genome of *E. meningoseptica,* finally obtaining 4 gapped fragments of *E. meningoseptica* genome. Mate Pair sequenced genomes were further analysed and annotated, showing different resistance patterns of each bacterium. Relation to other bacterial strains were further determined building a phylogenetic tree. Nevertheless, technological progress made HTS a powerful scientific tool, which can perform fast and efficient genome finishing thereby minimizing the effort and also providing an exceptional quality of the data.

Acknowledgements

I would like to acknowledge Dr Nick Loman for guidance and support throughout the project. I also would like to thank Josh Quick for help with sequencing and computational analysis as well as Calum Thomson for help in the laboratory.

Contents

Chart list

List of figures

Abbreviations:

ATM	Amplicon Tagment Mix
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
DAL	Diluted Amplicon Library
DMSO	Dimhethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide Triphosphates
E. coli, Ec	*Escherichia coli*
E. meningoseptica, Em	*Elizabethkingia meningoseptica*
FC	Final concentration
FU	Fluorescence intensity
GC content	Guanine - cytosine content
H_2O	Water
HAF	Hash Frequency
HS	High sensitivity
HT1	Hybridization buffer
HTS	High-throughput sequencing
LNA1	Library Normalization Additives 1
LNB1	Library Normalization Beads 1
LNS	Library Normalization Storage Buffer 1
LNW1	Library Normalization Wash 1
LR	Long range
MAR	Multiple Antibiotic Resistance
MIRA	Mimicking Intelligent Read Assembly
MP	Mate Pair
NaOH	Sodium hydroxide
NPM	Nextera PCR Master Mix
NT	Neutralize Tagment Buffer
ORFs	Open reading frames
P. aeruginosa, Pa	*Pseudomonas aeruginosa*
PAL	Pooled Amplicon Library
PCR	Polymerase chain reaction

PR2	Incorporation Buffer
RAST	Rapid Annotation using Subsystem Technology
Rep	Repetitive contig
RNA	Ribonucleic acid
RSB	Resuspension Buffer
SBS	Sequencing by synthesis
SNPs	Single nucleotide polymorphisms
SSPACE	Stand-alone scaffolder of pre-assembled contigs
TBE buffer	Tris-Borate-EDTA buffer
TD	Tagment DNA Buffer

1. Introduction

For many centuries microorganism-induced infectious diseases remained one of the most common causes of human illness and death. After invention and successful development of antibiotics it was assumed that all infectious diseases as well as disease experts will soon disappear. However, bacterial populations exposed to antibiotic stress demonstrated the opposite, starting a new era of infectious diseases during past 30 years. Bacterial organisms were able to effectively adapt in their environment and develop a resistance against drugs. Many strains became multidrug resistant, which is a rising concern nowadays as this can lead to panresistance and complete futility of medications. Currently the most challenging bacterial pathogens with developed drug resistance include numerous gram-positive bacteria (methicillin-resistant *Staphylococcus aureus*, multidrug-resistant *Streptococcus pneumoniae*, vancomycin- resistant *Enterococcus* spp.) as well as various gram-negative bacteria such as multidrug-resistant *Acinetobacter baumannii*, *Klebsiella pneumoniae*, *Escherichia coli*, *Pseudomonas aeruginosa* and *Elizabethkingia meningoseptica* (Lister et al., 2009).

1.1. *Elizabethkingia meningoseptica*

1.1.1. Biology and distribution

The genus of Elizabethkingia includes aerobic, Gram-negative, rod-shaped bacteria, which are non-motile and produce yellow coloration during the culture (rarely can be non-pigmented). Until 2011 only two species were representing the genus – *Elizabethkingia meningoseptica* and *Elizabethkingia miricola* (Bernardet et al., 2006), however in 2011 a new species – *Elizabethkingia anophelis* was discovered in the midgut of the mosquito *Anopheles gambiae* (Kämpfer et al., 2011). Both *E. meningoseptica* and *E. miricola* are able to cause human

infections, however *E. meningoseptica* is more potent pathogen and more often is associated with severe diseases. *E. meningoseptica*, earlier known as *Flavobacterium meningosepticum* and *Chryseobacterium meningosepticum*, is widely spread in the environment, especially in soil and water (Steinberg & Burd, 2010). Similar to *P. aeruginosa*, *E. meningoseptica* often infects immuno-compromised and postoperative patients as well as infants (Jacobs & Chenia, 2011).

The biotopes, inhabited by *E. meningoseptica* vary, but mostly have high humidity level. Those are different water bodies such as lakes and other freshwater, soil, nuclear fuel pools or even space (Bernardet et al., 2006). Bacterium was also isolated from various eukaryotic organisms, for instance, protozoa, fish, amphibians, reptiles, birds and mammals. In medical units *E. meningoseptica* is often isolated from many sources, including freshwater and saline solutions, disinfectants, medical equipment and devices, such as feeding tubes, respirators and catheters (Ceyhan et al., 2008).

1.1.2. Clinical picture

E. meningoseptica can be a part of normal microflora in different organisms, for instance, fish, however it cannot inhabit the microbial flora of a healthy human organism. Colonized individuals may not show observable symptoms; still these patients can potentially infect other susceptible organisms. Moreover, without appropriate therapy colonization will progress to clinical infection. *E. meningoseptica* most often infects individuals, who have malignant tumors, neutropaenia, diabetes, end-stage hepatic and renal disease, extensive burns or undergo dialysis, malnutrition and use steroids. *E. meningoseptica* is a common cause of nosocomial infections such as meningitis, sepsis, bacteremia, pneumonia, endocarditis, skin and soft tissue infections, wound infection, abdominal abscess, ocular infections, sinusitis,

bronchitis, epididymitis, dialysis-associated peritonitis and prosthesis-associated septic arthritis (Bloch et al., 1997; Ceyhan & Celik, 2011).

- **Neonatal meningitis**. The most common disease induced by *E. meningoseptica* is neonatal meningitis, which is especially frequent in children during their first weeks of life. The mortality rate in meningitis patients is especially high, reaching up to 57%. Moreover, after infection several postinfectious abnormalities are detected, including brain abscesses, hydrocephalus, deafness and developmental delay.
- **Bacteremia**. Second most common disease associated with *E. meningoseptica* is bacteremia, where bacteria get into the blood circulatory system. Bacteremia may arise as a complication of another disease (meningitis), after surgery or due to foreign bodies (catheters) in blood vessels. As a consequence it may initiate another infection, for instance, sepsis (Ceyhan & Celik, 2011).
- **Pneumonia** is another *E. meningoseptica* caused disease, occurring mainly in premature infants. Similar to bacteremia it is rapidly transmitted and frequently occurs as outbreak (Hoque et al., 2001).

Several infections, such as cellulitis, septic arthritis, community-acquired respiratory tract infection, keratitis and bacteremia were detected not only in immune-compromised patients, but also in healthy individuals (Steinberg, 2000).

1.1.3. Antibiotic resistance

Throughout the history bacterial ability to adapt to different environmental changes and resist various medications has been a serious challenge for scientists. Elizabethkingia are ubiquitous, multidrug resistant pathogens, therefore the right choice of antibiotics for the therapy is crucial. Poor treatment outcome and as a result the progression of infection usually

arise from the use of inactive or inappropriate bacterial agents against strains, which have highly specific resistance. Patterns and mechanisms of *E. meningoseptica* resistance are quite unusual and still poorly investigated. It is known so far that *E. meningoseptica* is resistant to various types of drugs, and particularly to β-lactams as resistance to β-lactams is chromosome-encoded. It produces 2 classes of β-lactamases – A or extended-spectrum β-lactamases and B or metallo-β-lactamases. The second group is responsible for the resistance to carbapenems that are extensively employed in treatment of infections induced by multidrug resistant Gram-negative microorganisms (Ceyhan & Celik, 2011). *E. meningoseptica* is the only bacterium, which has two types of metallo-β-lactamases – BlaB and GOB (Bellais et al., 2000).

In many cases *E. meningoseptica* is resistant to certain antimicrobial agents, which are employed in meningitis therapy, including polymyxins, aminoglycosides (such as gentamicin, streptomycin) penicillin, ampicillin and chloramphenicol. Bacterium is also resistant to tetracyclines, erythromycin, linezolid, clindamycin and vancomycin (several strains are intermediately susceptible to the last two). *E. meningoseptica* appears to be the most sensitive to the following drugs: minocycline, rifampicin and quinolones (Krieg, 2006).

1.1.4. Genetics & genomics

Information about the genetics and genomics of *E. meningoseptica* is very scarce. Some studies were conducted on DNA-DNA hybridization to determine the relatedness of *E. meningoseptica* to other species. GC content of the bacteria was also determined (being about 37%; Krieg, 2006) and several reports were investigating functions and properties of single genes (such as *BlaB* and *GOB*, which are responsible for the resistance to β-lactams; *gyrB*, which is also responsible for the resistance). All these investigations were accomplished for phylogenetic and taxonomic purposes or for the molecular typing of particular strains.

However, any further information about the genome structure, size, genes (their function and location) within the genome is not available. Nevertheless, genetic manipulation structures established for *Flavobacterium johnsoniae* including the gene transfer system, selectable marker system, suicide vector system and transposons were also successfully applied on *E. meningoseptica* (Dworkin et al., 2006).

It is important to note that no plasmid has been discovered in genera of Elizabethkingia (*Holmes & Owen, 1981*). Therefore, it can be concluded that the resistance to various antibiotics is most likely chromosome-encoded (Dworkin et al., 2006).

1.2. *Pseudomonas aeruginosa*

1.2.1. Biology and distribution

Pseudomonas genus mostly includes environmental saprotrophs, and the only species that can induce severe human diseases is *Pseudomonas aeruginosa* (Mathee et al., 2008). *P. aeruginosa* is a gram-negative, rod-shaped bacterium well known for its blue-green coloration created while cultured (Lister et al., 2009). It is asporogenous, monoflagellated microorganism, which demonstrates high nutritional flexibility and adaptability (Lederberg et al., 2000). *P. aeruginosa* is considered as aerobic organism, however can be anaerobic in the presence of suitable substrate. This species can live in various environments such as soil, marshes and coastal marine habitats, as well as be present on plant and animal tissues (Hardalo & Edberg, 1997). It can also form biofilms in humid biotopes, for instance, rocks and soil (Costerton et al., 1999).

The ability of *P. aeruginosa* to catabolize a variety of organic complexes, for instance, bacteriostatic benzoate makes this bacterium ubiquitous and highly persistent (Lederberg et al., 2000). As all opportunistic pathogens it rarely causes diseases in healthy organisms,

however individuals with compromised immune system are very likely to develop the infection. It is also one of the most common opportunistic pathogens in hospitals, responsible for a quite high mortality (40 – 60%; Fick, 1993). In medical care units *P. aeruginosa* is often found on therapy equipment, antiseptics, soap, sinks, mops, medicines, as well as in physiotherapy and hydrotherapy pools. Apart from hospitals is can be also present in public swimming pools, whirlpools, hot tubs, contact lens solution, home humidifiers, soil, rhizosphere and various food products including vegetables (Pollack, 1995).

1.2.2. Clinical picture

Pseudomonas aeruginosa very rarely is a part of the normal microflora in human organism. It can be present in different tissues at a small percentage (0-6.6%); however during hospital treatments colonization may be even higher than 50% (Pollack, 1995). Patients with trauma, for instance, burns, or immunocompromised individuals (with cystic fibrosis, cancer) have higher risks to be infected by *Pseudomonas.* In spite of high abundance of this bacterium in the environment almost all severe infections associated with *P. aeruginosa* are predominantly acquired during hospitalization (Lister et al., 2009). Previous studies indicated that *P. aeruginosa* is the second most common pathogen isolated from the intensive care unit patients in Europe. It is identified approximately in 30% of pneumonia cases, 19% of urinary tract infection and 10% of bloodstream infection cases (Spencer, 1996).

P. aeruginosa can target different organs and tissues of the human body and subsequently cause various types of infection.

- **Eye infections**: Involve the damage of cornea mostly by contact lenses, and bacteria get into the eye as a result of injury.

- **Ear infections**: Outer ear canal infection occurs when water remains in the ear after swimming. It can progress to malignant otitis externa, invading deeper structures of the ear.
- **Chronic respiratory infections**: Predominantly develop in patients with pulmonary diseases such as cystic fibrosis and lead to rapid decrease in lung function. Lung infections can also develop in case of bronchiectasis and obstructive pulmonary disease.
- **Hospital-acquired pneumonia**: Widely spread disease, which in many cases is associated with *P. aeruginosa*, particularly in mechanically ventilated patients. It is characterized by a high number of deaths.
- **Urinary tract infections**: One of the most common types of infection caused by *P. aeruginosa.* Can develop after surgery or due to a foreign body in urinary tract.
- **Bloodstream infections**: Characterized by nodular skin lesions, and caused mainly by *P. aeruginosa.*
- **Skin and soft tissue infections**: Can be acquired in hot tubs through softened skin and lead to folliculitis. Patients with severe burns are at a particular risk of infection as *P. aeruginosa* very often colonizes wounded skin (Hauser & Ozer, 2011)

1.2.3. Antibiotic resistance

P. aeruginosa is a highly resistant, problematic pathogen acquired in community or hospitals (Micek et al., 2005). Due to *P. aeruginosa* ability to resist numerous types of drugs as well as to create resistance even during the therapy, it is difficult to choose appropriate medications. Diseases induced by these multidrug resistant bacteria are associated with high morbidity and mortality rates, often involve surgeries, prolong hospital stay and chronic care, and increase

the cost of treatment (Aloush et al., 2006). Even more complicated clinical cases arise if the bacterial resistance is formed during the therapy (Dimatatac et al., 2003).

P. aeruginosa can develop multidrug resistance through several pathways:

- Through the obtaining of multiple imported resistance on mobile genetic components (for instance, plasmids),
- Through the mix of imported and chromosome encoded resistance,
- Through multiple modifications in chromosomal material over time
- And/or through a single mutation, which induces an overexpression of a multidrug resistance mechanism (for example, efflux pump).

Each resistance mechanism influences different types of drugs therefore can considerably restrain antibiotic selection possibilities and complicate the therapy of severe infections. Thus, for instance, imported resistance affects β-lactams and aminoglycosides, but not fluoroquinolones. In the case of β-lactams *P. aeruginosa* produces β-lactamases (for example, penicillinases) that inactivate β-lactams and consequently limit efficacy of the drug. On the contrary chromosome encoded resistance influences mainly fluoroquinolones (Lister et al., 2009). As well as imported resistance chromosome encoded resistance mechanisms vary. These can include aminoglycoside- inactivating enzymes (Poole, 2005) and D oxacillinases (Girlich et al., 2004).

The external membrane of gram-negative bacterium also contributes to its resistance. The membrane serves as a semipermeable barrier that slows down the diffusion of many antibacterial agents. In comparison to *E. coli* the external membrane of *P. aeruginosa* is much harder to penetrate. Nevertheless, to stay alive, it is important for the bacterium to receive nutrients, what is achieved using a group of water-filled protein channels – porins. These porins are key elements in the transport of sugars, amino acids, phosphates, divalent cations, and siderophores (Hancock & Brinkman, 2002). The loss of certain porin, however, initiates a

decline in membrane permeability, consequently decreasing effectiveness of antibiotics. Besides the loss of porins, decrease in drug accumulation is also accomplished actively exporting antibiotics from the cell by membrane-associated pumps – efflux pumps. All mentioned resistance strategies and, moreover, the ability to coregulate resistance mechanisms makes *P. aeruginosa* versatile and hardly curable pathogen, which still challenges therapeutic schemes (Lister et al., 2009).

1.2.4. Genetics & genomics

The genome of *P. aeruginosa* is noticeably larger than many other bacterial genomes sequenced so far. It consists of a single circular chromosome – 6.3 million base pairs (Mbp) for PAO1 strain (Stover et al., 2000) and with approximately 66% GC content. There are 5,570 predicted open reading frames (ORFs) within PAO1 genome, and it is almost as complex as the genome of the primitive eukaryote *Saccharomyces cerevisiae* (Ball et al., 2000). The majority of *P. aeruginosa* gene encoded products can be allocated into 3 functional groups: putative enzymes, transcriptional regulators and transporters of small molecules.

Amongst all bacterial genomes sequenced, *P. aeruginosa* genome appears to have the largest amount of predicted regulatory genes. Previous study identified about 468 genes which have similar characteristics to transcriptional regulators or environmental sensors (Stover et al., 2000). Following these results it can be concluded that around 8.4% of *P. aeruginosa* genes participate in regulation what is far more than in any other bacterial species. Moreover, the overexpression of transcriptional regulators such as two-component regulatory system proteins or sensor-response regulator hybrids allows the organism rapidly adapt to the changing environment.

The genes that encode outer membrane proteins are very important as they play a special role in antibiotic transport and export of extracellular virulence factors, as well as in anchoring the structures that facilitate adhesion and mobility of the cell. In *P. aeruginosa* the number of predicted genes that encode outer membrane proteins is particularly large (150, predicted by Stover et al., 2000). Furthermore, in cytoplasmic membrane *P. aeruginosa* has a variety of membrane transport systems, from which more than a half are responsible for the import of nutrients and other compounds.

Intrinsic resistance of *P. aeruginosa* to a range of drugs arises mostly from its external membrane permeability and active efflux of antibiotics (Hancock, 1998). Previous studies indicated that *P. aeruginosa* has four multidrug efflux systems, and all of them are a part of the resistance-nodulation-cell division family (Westbrock-Wadman et al., 1999; Nikaido, 1998). Furthermore, *P. aeruginosa* was reported to have the most intricate chemosensory system in comparison to other bacteria. It was found that there are 4 loci, which most likely encode signal transduction in chemotaxis. First locus in its structure resembles the locus of *Salmonella typhimurium*, which is responsible for sensoring chemoattractants (Stock & Surette, 1996), while second is similar to the locus of *Rhodobacter sphaeroides* (Armitage & Schmitt, 1997). The other two loci have similarities with the genes of *E. coli* and gliding bacterium, *Myxococcus xanthus* (McBride et al., 1989; Stover et al., 2000).

1.3. *Escherichia coli*

Escherichia coli is a Gram-negative rod-shaped bacterium and the most intensively studied prokaryote, used as a model organism in various scientific areas. Bacterium is a part of normal gut microflora in endothermic organisms including humans. However, several serotypes are serious human pathogens, often associated with food poisoning and contamination; these include enteropathogenic, enteroaggregative, enterotoxigenic,

enteroinvasive and Shiga toxin-producing *E. coli*. These strains are usually multidrug resistant and can cause outbreaks of severe infections including uremic syndrome (Suma et al., 2011) and meningitis (Lu et al., 2011) as well as can complicate the disease such as inflammatory bowel disease (Crohn's disease; Miquel et al., 2010). One of the examples of multidrug-resistant *E. coli* includes the most recent outbreak in Germany caused by O104:H4 strain in 2011. This Shiga-toxin producing strain (genome size – 5.44 Mb) induced hemolytic–uremic syndrome, described by the acute renal failure, haemolytic anaemia, thrombocytopenia and bloody diarrhoea (Frank et al., 2011). The strain was characterised as enteroaggregative, Shiga toxin-producing and contained a plasmid encoding extended-spectrum β-lactamase (Grad et al., 2012) and resistance to cephalosporins and fluoroquinolones (Frank et al., 2011).

In current study *E. coli* (O104:H4 strain) will be used in Nextera optimization step as organism with average GC and in subsequent steps (Mate Pair sequencing and annotation) as a control organism.

1.4. Genome sequencing

1.4.1. High-throughput sequencing

Being a very common technique nowadays, sequencing allows user to decode the DNA sequence of a targeted organism. Following capillary sequencing methods, which rely on Sanger sequencing (Sanger et al, 1977), Next-generation or High-throughput sequencing (HTS) technology was recently developed. Unlike Sanger sequencing, which uses chain-termination principle, High-throughput sequencing employs highly parallelized processes, permitting to sequence thousands or millions of molecules at the same time. The most commonly used types of HTS are 454 pyrosequencing (Life Sciences, now Roche), which involves the detection of single nucleotides incorporated into DNA template employing

luciferase, SOLiD sequencing (Life Technologies), which employs special ligation of constant-length oligonucleotides and Illumina sequencing (Soon et al., 2013).

The SBS (sequencing by synthesis) technique, designed by 'Manteia Predictive Medicine S.A.' in late 1996 and owned by Illumina from 2006, is one of the most widespread and effective next-generation sequencing methods. TruSeq sequencing employs highly parallel sequencing based on reversible terminator method that allows recognition of single bases as soon as they incorporate into developing DNA strands. Each time when dNTP is incorporated the fluorescent terminator is imaged and further removed to enable another base integration. The data of this base-by-base sequencing is highly precise and has a high potential to be applied in many other research areas.

1.4.2. Whole-genome sequencing advantages and applications

The greatest benefit of whole-genome sequencing is in yielding the entire available DNA information of an isolate in just one simple step after culture. As a result all the data required for diagnostic and typing will be obtained, however the interpretation of this data is not always obvious. Genome sequence also contains a huge amount of additional information that cannot be obtained with routine techniques, offering the perspective of large-scale studies into bacterial genotype–phenotype relations.

In clinical microbiology whole-genome sequencing can be even more beneficial. The introduction of sequencing into clinical microbiology is notably simplified by genetic characteristics of bacterial organisms. In comparison to eukaryotes, bacterial genomes are quite small (2-6 Mb), and mostly represented by one haploid chromosome. Nevertheless, they are more diverse than the genomes of eukaryotes, as may contain dispensable elements (for instance, plasmids), which are not present in all members of the species (Didelot et al., 2012).

Theoretically, the genome sequence of a pathogen holds almost all information, which is essential to develop the treatment and establish public health measures. Moreover, fast and low-cost sequencing of a whole genome will most likely replace complex multistep techniques, which are now employed in characterization of isolated pathogens (Relman, 2011). However, it is still a challenge for the modern microbiology, which can be overcome only if the genomic knowledge and analytical techniques will allow extracting and interpreting the existing data appropriately. Furthermore, for successful technique introduction into clinical microbiology as a routine method, special attention should be paid when validating genotypic prediction of phenotype and especially for bacterial resistance studies (Didelot et al., 2012).

Whole-genome sequencing can be used in various research areas:

- **Species identification** is a very important stage of infection management and pathogen tracking. Taxonomic methods employed now rely on using the collection of different strain kinds as a standard, what in whole-genome sequencing can be substituted by "reference sequence", which is compared to isolate sequence.
- **Drug resistance prediction**. Using whole-genome sequencing tools the prediction of resistance phenotype (by the identification of resistance genes) should be possible. This will allow rapid selection of antibiotic therapy, however genetic determinants of drug resistance should be constant across the phenotypes (successful examples include *S. aureus* (McAdam et al., 2011), *Vibrio cholera* (Mutreja et al., 2011) and *Burkholderia dolosa* (Lieberman et al., 2011)). In combination with extensively used PCR techniques whole-genome sequencing can improve the resistance prediction results. Moreover, computational approach may be more sensitive than use of PCR primers (Didelot et al., 2012).

- **Identification of virulence factors**. It is essential to improve the knowledge of genetic mechanisms of virulence, as they are extensively studied, but not well-understood. Sequencing can provide all of the information about virulence factors in a single step, also opening a horizon for new virulence factor discovery employing the connection of isolate data with outcome data from the patient (example includes identification of Shiga toxin in O104:H4 strain of *E. coli;* Rasko et al., 2011).
- **Outbreak detection and surveillance**. Sequencing has a potential to deliver well-resolved, precise, high-quality data regarding isolate relatedness. Whole-genome sequencing of various organisms (for instance, *Chlamydia trachomatis;* Harris et al., 2012) already revealed limitations of existing clinical typing methods, which are used for identifying phylogenetic relationships. Latest studies also indicated that sequencing can be very useful in outbreak source detection and in transmission route identification within societies and medical units (Koser et al., 2012; Eyre et al., 2012; Didelot et al., 2012).

1.4.3. Fragments, pairs and mate-pairs

Basic sequencing method, which relies on chain-termination principle, is still widely used nowadays. However, it can decode only short DNA fragments (1,000-2,000 base pairs (bp)), which is an important limitation as the most simple viral genome already consists of more than 10,000 bp, bacterial genomes are even larger – several millions bp, and mammalian genomes consist of billions of base pairs. A relatively new principle – shotgun sequencing was developed as a solution for this problem. The principle is based on shearing the DNA sequence into multiple short fragments, followed by sequencing of fragment ends. Produced sequences are then joined together or assembled by a computer program – assembler (CBCB, cited June 2013).

The underlying principle of the assembler is based on assumption that two sequence reads, which contain same sequence of nucleotides, have the same location in the genome. The data employed in genome assembling consist of single and paired reads. Single reads are originally sequenced small DNA fragments. Overlapping these sequences the programme can join them together, creating a contiguous (ungapped) sequence or contig (Baker, 2012); however it can create an assembly only if sequences are able to cover the genome 8-10 times. This oversampling of the genome is known as coverage. A perfect assembly program should generate only one contig for every sequenced chromosome. Nevertheless, in almost all cases a number of contigs are created due to unsequenced parts of the genome and occurrence of multiple almost-indistinguishable copies of genome regions so called repeats, which can cause errors (CBCB, cited June 2013).

Paired reads have similar sizes to single reads, however, they are originated from either 5' or 3' end of DNA fragments, which are too large for straight-through sequencing (Baker, 2012). Paired-end sequencing allows obtaining reads from DNA fragments in both forward and reverse directions. If the additional information, for instance mate-pairs, is available, generated contigs can be ordered and oriented along a chromosome, producing scaffolds (ordered contigs with estimated gaps in between). Mate-pair fragments are much larger (2-5 kb) than fragments obtained with original paired-end sequencing and contain junction adapter, therefore permitting estimation of gaps between contigs and orientation of contigs (Nagarajan et al., 2010).

1.4.4. *De novo* assembly

As the outcome of high-throughput sequencing is a variety of short (30 to 100 bp) overlapping reads from the desired genome, the initial step of the analysis is assembling the reads into longer genome fragments (Nielsen et al., 2011). There are two methods to

accomplish that: to create a 'reference based assembly' or to produce *de novo* assembly. The first approach compares the reads to an available reference genome (sequence of which is already determined) to see which regions are alike and which different. This method is very efficient, when determining genetic diversity in isolates with available high-quality reference genome. Disadvantages of such method are: any fragment, which is not present in the reference genome, will not be assembled; mapping accuracy decreases, if the target and reference genomes are not closely related. Therefore, genome assembling without the reference or *de novo* assembly (carried out by programs Newbler (Chaisson & Pevzner, 2008) or Velvet (Zerbino & Birney, 2008)) is of a particular interest nowadays. *De novo* approach does not possess previously mentioned problems: the whole genome is assembled and quality is not reference-dependant (Didelot et al., 2012). Moreover, *de novo* assembly can provide more information, when sequencing a new organism or a new strain of well-known species (Loman et al., 2012). *De novo* assembly, however, is complicated by its result, which gives tens or hundreds of contigs. Due to the complexity of the target genome and the presence of repetitive fragments, which can be homologous to many genome regions, further assembling can become even more difficult or almost impossible (Didelot et al., 2012).

1.4.5. Genome finishing

Prior the invention of HTS, it took several weeks to obtain a draft sequence of a particular genome, while to finish the whole genome (obtain a complete sequence) many month or even years of intensive (and expensive) experiments were required. Development of HTS started a new era of genome finishing, where a single machine can produce draft assemblies for 5 bacterial genomes in just few days. Finished or nearly-finished genome is a desirable outcome that allows a scientist to acquire broader and richer information for genomic analysis. To

simplify the finishing process draft assemblies can be combined with additional sequence and map-based information (Nagarajan et al., 2010).

There are two main limitations of shotgun sequencing – genome fragmentation (due to repeats) and errors (due to sequencing artefacts or improper reconstruction of repeats), which finishing process aims to conquer. Finishing, therefore, consists of two main phases: gap closure, and assembly validation and enhancement. First phase includes determination of the sequence, which fits into the gap between adjacent contigs employing directed-PCR and primer-walking approaches. With the presence of mate-pair data contig linkage can be quite easily determined and gaps can be closed. However, contigs, for which mate-pairs cannot provide linkage information, require time-consuming, expensive and various-combination PCR (Tettelin et al., 1999). The second finishing phase – validation and enhancement – targets to correct errors in the assembly. These errors can be single-base, for instance, if a base is mis-called, and large-scale, for example, mis-assemblies due to repetitive regions. Nevertheless, errors caused by repeats can be fixed using targeted PCR aimed to correct the adjacency of the assembled contigs (Nagarajan et al., 2010).

1.5. Aim and objectives

This study **aims** to finish or start finishing the genomes of *Pseudomonas aeruginosa* and *Elizabethkingia meningoseptica* as well as to optimize Nextera protocols for bacteria with different GC.

The **objectives** are:

- To optimize Nextera XT protocol for bacteria with different GC content using *Elizabethkingia meningoseptica*, *Escherichia coli* and *Pseudomonas aeruginosa*, and obtain 450 bp or longer fragments;
- To join sequencing obtained fragments into contigs;
- To perform Mate-pair sequencing employing *Elizabethkingia meningoseptica*, *Escherichia coli*, and *Pseudomonas aeruginosa* 901 and 950 and yield long sequences (2-5 kb);
- To join contigs into scaffolds;
- To finish the genome employing PCR and Sanger sequencing;
- To annotate the genomes of bacteria.

2. Materials and Methods

2.1. Sample preparation for Nextera testing

2.1.1. Isolates

Elizabethkingia meningoseptica strain 501 was isolated in Birmingham Hospital from blood of 60 year old patient with acute leukaemia. Susceptibility tests showed that specimen is resistant to piperacillin/tazobactam, ceftazidime, meropenem and aztreonam, while sensitive to amikacin and ciprofloxacin, and variable to gentamicin. *Pseudomonas aeruginosa* ST395 specimen 901 was isolated from water samples (in hospital shower) and specimen 905 was sampled besides *Pseudomonas*-positive patient in Birmingham hospital. *Escherichia coli* strain O104:H4 specimen 1218_280 was received from London Health Protection agency (extracted from patient, who arrived to UK from Germany, carrying Shiga toxin-producing *E. coli*). DNA from *Elizabethkingia meningoseptica* 501 and *Pseudomonas aeruginosa* ST395 was extracted in Birmingham Hospital.

2.1.2. Parameter optimization for bacteria with different GC

In order to optimize Nextera protocols for bacteria with different GC content 3 bacteria were involved in the experiment: *Pseudomonas aeruginosa* ST395 with high GC content (~65%), *E. coli* O104:H4 with medium GC (~50%) and *Elizabethkingia meningoseptica* 501 with low GC (~38%). A stock solution of 4 ng/µl was firstly obtained diluting the original sample in elution buffer (EB; QIAGEN). Further concentrations were obtained diluting the appropriate amount of stock solution in EB. Four concentrations of input DNA were used: 0.1, 0.2, 0.3, and 0.4 ng/µl, for every species. Reaction times in thermocycler during tagmentation process were also optimized: 3, 5 and 7 minutes.

2.2. Fluorometer

To determine original and diluted sample concentrations high sensitivity (HS) dsDNA assay was carried out. High sensitivity Quant-iT™ dsDNA assay kit (Molecular Probes, Life technologies) was used to perform precise determination of concentrations. After kit reagents reached room temperature, working solution was prepared by diluting HS reagent in HS buffer (1:200). It was transferred to the Qubit assay tubes (Molecular Probes, Life technologies), both standard and sample tubes. Standard 1 (λ dsDNA High sensitivity standard 0 ng/ µl) and 2 (λ dsDNA High sensitivity standard 10 ng/ µl) were then added to standard tubes (1:20 in working solution). After Qubit flourometer (Invitrogen) was programmed to run dsDNA HS assay, it was calibrated using standards. Next, samples were added to sample tubes (1:200 in working solution) and quantified measuring the fluorescence at excitation/emission 500/520.

2.3. Nextera XT sequencing

The Nextera XT Sample preparation kit allows preparing isolated DNA for sequencing on various Illumina sequencing machines. All reagents used in Nextera XT experiments were purchased from Illumina Inc., unless stated otherwise (Nextera XT protocol – Illumina, 2012a).

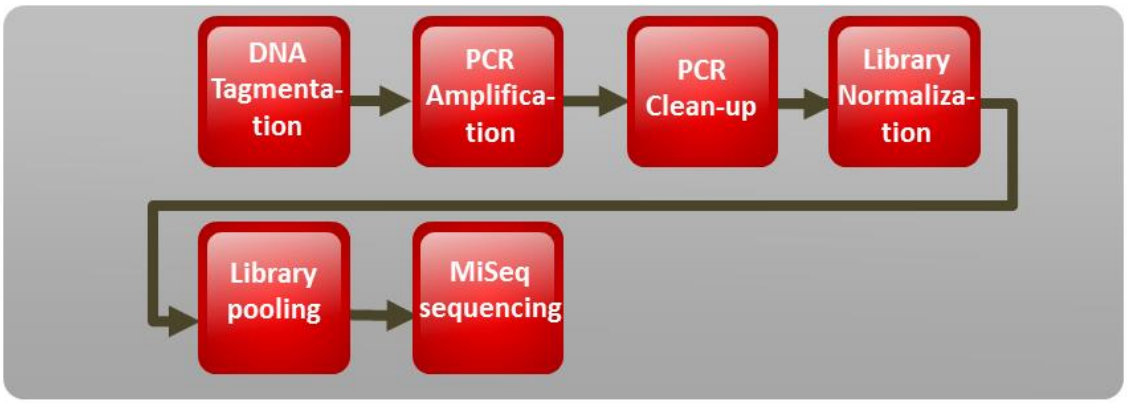


Figure.2.1. Nextera XT workflow. Adopted from Illumina protocol (Illumina, 2012a).

2.3.1. DNA tagmentation

First, all the reagents were thawed, vortexed (Vortex genie, Scientific Industries Inc.) and briefly centrifuged in a microcentrifuge (Eppendorf, 5417C). Then TD buffer (Tagment DNA Buffer) was added to all used wells of 96-well PCR plate (BIO-RAD Laboratories Inc.). Input DNA (0.1, 0.2, 0.3, and 0.4 ng/µl) was then transferred to the plate, and ATM (Amplicon Tagment Mix) was added to the mix. This solution was gently mixed, the plate was covered with Microseal (seal 'B', BIO-RAD Laboratories Inc.) and centrifuged (Eppendorf, Centrifuge 5810R) at 280 xg (x gravitation or rcf) for 1 minute. Next the plate was transferred to thermocycler (Eppendorf, Mastercycler Gradient) and following thermocycling conditions were set:

- 55 ºC for 3, 5 or 7 minutes
- Hold at 10 ºC.

After samples reached 10 ºC Microseal was removed and NT (Neutralize Tagment Buffer) was added. Then plate was covered again, centrifuged at 280 xg for 1 minute and incubated at room temperature for 5 minutes.

2.3.2. PCR amplification

NPM (Nextera PCR Master Mix) was added to each well of the plate from previous step. Then the unique combination of 2 index primers (one N7XX and one S5XX primer type) was added to each well. The solution was gently mixed, and the plate was sealed with Microseal (seal 'A', BIO-RAD Laboratories Inc.) and centrifuged at 280 xg for 1 minute. PCR amplification was then carried out using the following thermocycling conditions:

- 72 ºC for 3 minutes

- 95 ºC for 30 seconds
- 12 cycles of:
    - 95 ºC for 10 seconds
    - 55 ºC for 30 seconds
    - 72 ºC for 30 seconds
- 72 ºC for 5 minutes
- Hold at 10 ºC.

After amplification samples were either cleaned-up immediately or stored in thermocycler at 10 ºC overnight.

2.3.3. PCR clean-up

PCR clean-up was performed to purify DNA library and remove very short fragments from the samples. First, AMPure XP beads (Agencourt, Beckman Coulter Inc.; 0.5x) were added to PCR plate and mixed with the sample. The plate was incubated at room temperature for 5 minutes and placed on a magnetic stand (DynaMag™ - 96 Side skirted, Life Technologies AS) for 2 minutes. When the supernatant appeared to be clear it was removed from wells. Then beads were washed with 80% ethanol (diluted in distilled H_2O; Fisher Scientific), the plate was incubated at room temperature for 30 seconds and supernatant was discarded. Ethanol wash was then performed for the second time. While the plate still remained on the magnetic stand, beads were air-dried for 15 minutes. Next, the plate was removed from the stand and RSB (Resuspension Buffer) was added to each well. After mixing the solution was incubated at room temperature for 2 minutes. The plate was then placed on the magnetic stand for 2 minutes, and cleared supernatant was transferred to a new plate. After clean-up samples were either normalized immediately or stored at -25 ºC for 2-3 days.

2.3.4. Library normalization

Normalization process was carried out in order to ensure that all samples were equally represented in the library. In a separate 15 ml Falcon tube (BD Falcon™) LNA1 (Library Normalization Additives 1) was mixed with LNB1 (Library Normalization Beads 1) and this mix was then added to each well of the plate containing DNA from previous step. Plate was covered with Microseal B and shaken on a microplate shaker (Eppendorf, Thermomixer comfort, Hamburg, Germany) at 1,800 rpm for 30 minutes. Next, the plate was positioned on the magnetic stand for 2 minutes. After supernatant was clear, it was removed. Plate was removed from the stand and washed with LNW1 (Library Normalization Wash 1) twice. NaOH 0.1N (10M Sodium hydroxide solution, Sigma-Aldrich; diluted 1:100) was then added for sample elution. Plate was then placed on the magnetic stand for 2 minutes. LNS1 (Library Normalization Storage Buffer 1) was then added to a new storage plate and clear supernatant was transferred from the previous plate. The plate now containing normalized samples was sealed and centrifuged at 1,000 xg for 1 minute. The plate was stored at -25 ºC or pooled straightaway.

2.3.5. Library pooling

The heat block (Eppendorf, Thermomixer comfort, 0 rpm) was set to 96 ºC and MiSeq reagent cartridge was thawed at room temperature. Normalized plate was centrifuged at 1,000 xg for 1 minute and samples were mixed. The libraries from each well were combined in one Eppendorf tube (1.7 ml, Sarstedt) labelled PAL (Pooled Amplicon Library) and mixed. The content of PAL was then transferred into another tube (DAL, Diluted Amplicon Library) containing HT1 (Hybridization buffer) and mixed again. After DAL was incubated at 96 ºC for 2 minutes, inverted and immediately placed in ice-water bath for 5 minutes. Finally, DAL was loaded into MiSeq reagent cartridge (Well 17 - Load Sample reservoir).

2.3.6. MiSeq Sequencing

After loading Amplicon Library into MiSeq reagent cartridge, sequencing run was set, following the directions on the screen. A flow cell (Single-use PE flow cell) provided in the kit was washed in dH_2O and wiped with methanol (Absolute, Fisher Scientific) prior loading into Flow Cell Stage. The reagents were loaded into Reagent Compartment: first, the PR2 bottle (Incorporation Buffer) and empty waste bottle, and after the reagent cartridge. When the software confirmed that all components and conditions were correctly set, sequencing (24 hours) was performed.

2.4. Bioanalyzer

2.4.1. High Sensitivity assay for Nextera XT

To perform a quality control of DNA fragment sizes High Sensitivity DNA assay was carried out employing Agilent 2100 Bioanalyzer on different sequencing stages (after PCR amplification and PCR clean-up). The electrodes of bioanalyzer were washed with dH_2O prior the run. High Sensitivity DNA Kit (Agilent Technologies) was used in the assay. First, all reagents were equilibrated to room temperature (30 minutes) and vortexed. Then, HS DNA dye concentrate was added to gel matrix and mixed well. The complete mix was transferred to spin filter tube, vortexed until appeared homogenous and centrifuged (Mikro 22 Centrifuge, Hettich Zentrifugen) for 10 minutes at 6000 rpm. The filter was removed, and the gel-dye was used in subsequent steps.

High Sensitivity DNA chip (Agilent Technologies) was placed on the chip priming station (Agilent Technologies), and gel-dye was added to the well marked G. After pressurising, gel-dye was added to 3 wells marked G. HS DNA marker was inserted in all the remaining wells, ladder was added to ladder well, and samples were added to 11 remaining wells. DNA chip

was then vortexed (Vortexer Ika Works) for 1 minute at 2400 rpm and inserted into Bioanalyzer. High Sensitivity DNA assay was selected in 2100 Expert software, and the run was performed (Agilent Technologies, 2009).

2.4.2. DNA 12000 assay for Nextera Mate pair sequencing

Quality control of mate pair fragments was performed at different sequencing stages (after AMPure Purification of Strand Displacement Reaction and during validation step) using Agilent 2100 Bioanalyzer 12000 assay kit (Agilent Technologies). The gel-dye was prepared as in HS assay. 12000 DNA chip (Agilent Technologies) was used in the assay. Gel-dye was added to the well marked G, pressurized, and again added to 2 wells marked G. The marker was inserted in all the remaining wells, ladder was added to ladder well, and samples were added to 12 remaining wells. DNA chip was then inserted into Bioanalyzer. 12000 DNA assay was carried out using 2100 Expert software (Agilent Technologies, 2007).

2.5. Nextera Mate Pair sequencing

Nextera Mate Pair sample preparation was performed using gel-free protocol as it is shorter, more robust, yields a higher diversity of fragments, and requires less input DNA than the gel-plus protocol. All reagents were supplied by Illumina, unless stated otherwise (Nextera Mate Pair protocol – Illumina, 2013).

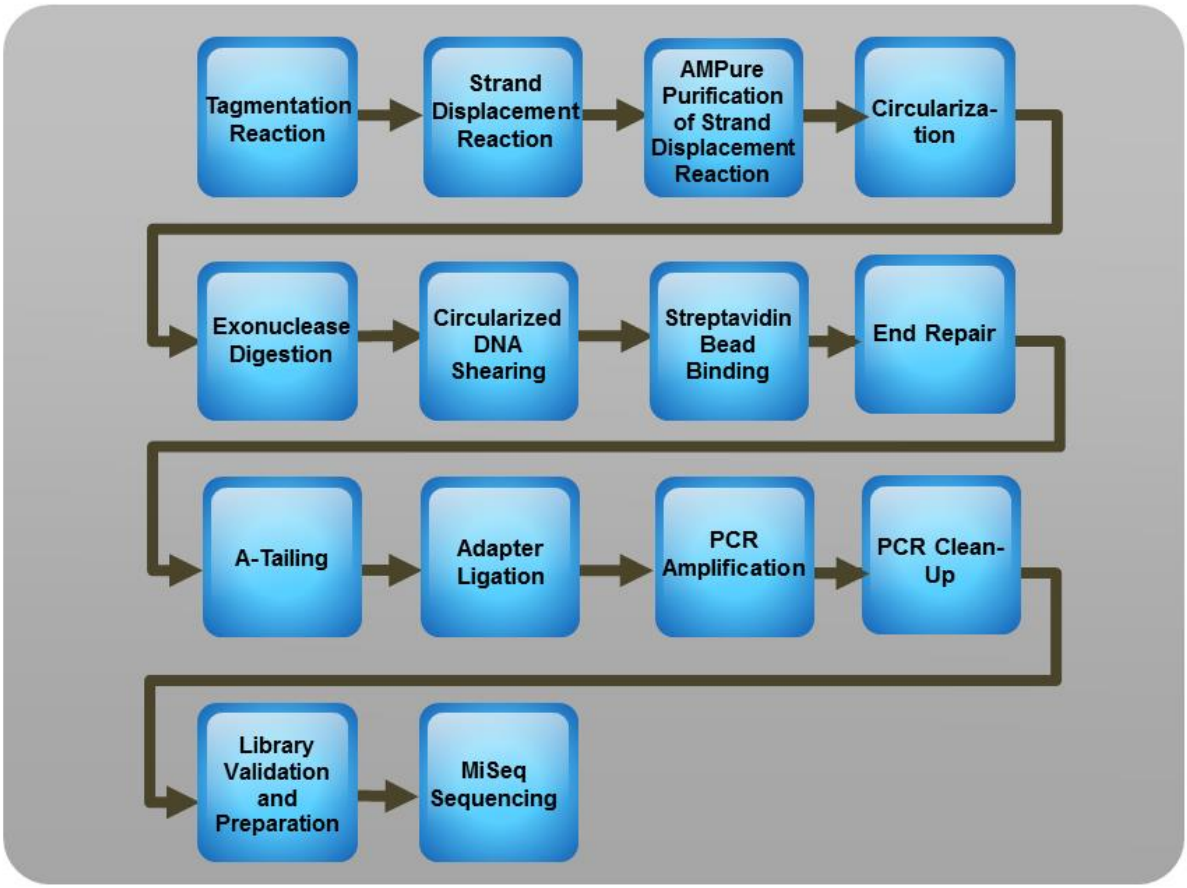


Figure.2.2. Nextera Mate Pairs workflow. Adopted from Illumina protocol (Illumina, 2013).

2.5.1. Tagmentation

First, the reaction components were added to the eppendorf tube in following order: Genomic DNA (1 µg), dH_2O, Mate Pair Tagment Buffer and Tagment Enzyme. The tube was incubated at 55 ºC for 30 minutes. Then tagmentation reaction was purified employing Zymo Genomic DNA Clean & Concentrator™ kit (Zymo Research). Briefly, ChIP DNA Binding Buffer (2 vol) was mixed with the samples, transferred to a Zymo-Spin IC-XL column in a collection tube, and mixture was centrifuged for 30 seconds at 11,000 xg. Wash step (repeated twice) was performed using Zymo DNA Wash Buffer (centrifugations were performed for 1 minute at 11,000 xg). The column was then transferred to a clean tube, and RSB (Resuspension Buffer) was added to elute DNA.

2.5.2. Strand Displacement

Strand displacement was carried out immediately after tagmentation. Reagents were added to the tube with tagmented DNA in the following order: water, 10X Strand Displacement buffer, dNTPs and Strand Displacement Polymerase. The reaction was mixed, briefly centrifuged (1,200 rcf, 3-5 seconds) to collect reaction to the bottom, and the tube was incubated at 20 ºC for 30 minutes.

2.5.3. AMPure Purification of Strand Displacement Reaction

The reaction components were added to the strand displaced DNA from previous step in the following order: water and then AMPure XP Beads. The reaction was incubated at room temperature for 15 minutes, and the tube (centrifuged at 1,200 rcf, 3-5 seconds) was placed on the magnetic stand for 5 minutes. Beads were then washed in 70% ethanol (7:3 in dH_2O) twice. After air-drying beads (15 minutes), RSB was added to elute DNA from beads. The tube was then briefly centrifuged, incubated at room temperature for 5 minutes and placed on magnetic stand for 5 minutes. The supernatant, containing purified DNA was transferred to a new tube. At this time point to assess the quality of the fragments, samples (diluted in dH_2O – 1:1) were examined using Bioanalyzer (1200 DNA assay).

2.5.4. Circularization

Circularization was carried out straight after the quality control (due to the gel-free protocol, size selection step was skipped). The following components were added to a clean eppendorf tube: purified DNA (500 ng), water, Circularization Buffer (10X) and Circularization Ligase. The reaction was mixed by inverting the tube and briefly centrifuged (1,200 rcf, 3-5 seconds). The tube was incubated at 30 ºC for 12-16 hours (overnight).

2.5.5. Exonuclease Digestion

After overnight incubation digestion of linear DNA was performed. Exonuclease was added to the tube with circularized DNA from previous step, and the tube was briefly centrifuged (1,200 rcf, 3-5 seconds). After it was transferred to the heat block and incubated at 37 ºC for 30 minutes. Then exonuclease was inactivated by incubation at 70 ºC for 30 minutes. Stop Ligation buffer was added to the tube, and the solution was mixed and briefly centrifuged.

2.5.6. Circularized DNA Shearing

After digestion DNA was shared into small fragments using Covaris AFA™ Ultrasonicator S2 (Covaris Inc.). Following the guidelines, the water in the instrument was de-gassed and pre-chilled to 3-6 ºC. The sample was transferred to Covaris glass tube, and the tube was capped. The subsequent parameters were applied to the Covaris programme: Intensity – 8, Duty Cycle/Duty Factor – 20%, Cycles per Burst – 200, Time – 40 s, Temperature – 6 ºC. After shearing process the sample was transferred to new eppendorf tube.

2.5.7. Streptavidin Bead Binding

Streptavidin bead binding process was performed immediately after shearing. Dynabeads M-280 streptavidin magnetic beads (Invitrogen) were placed (in a tube) on the magnetic stand for 1 minute and then washed (twice) and resuspended in bead bind buffer. The complete beads were added to sheared DNA sample and incubated at 20 ºC for 15 minutes. The **bead wash** was then performed (on the magnetic stand, 4 times), employing a bead wash buffer. After removing beads from magnet, they were resuspended in RSB and placed back on the magnetic stand for 30 seconds. The final solution remained on the stand until the next step.

2.5.8. End Repair

The following component was added to DNA (after supernatant removal) from previous step: end repair mix in water. The tube was then removed from magnet, and incubated at 30 ºC for 30 minutes. After samples were centrifuged and placed on the magnetic stand for 1 minute. The **bead wash** from Streptavidin bead binding step was then performed again.

2.5.9. A-Tailing

In a clean tube A-Tailing mix was combined with water. It was then added directly to the beads, and the tube was then transferred to the heat block and incubated at 37 ºC for 30 minutes.

2.5.10. Adaptor Ligation

After incubation adaptor ligation was carried out. The following ingredients were added to DNA from previous step: Ligation mix, water, DNA Adapter Index (AD0XX). After incubation (at 30 ºC for 10 minutes), Ligation Stop buffer was added to the reaction. The beads were briefly centrifuged and placed on the magnetic stand for 1 minute. The **bead wash** was then performed one more time.

2.5.11. PCR Amplification

The components of PCR reaction were added to the tube as follows: PCR Master Mix, PCR Primer Cocktail, water. After PCR reaction mix was added, and beads were transferred to PCR tubes (200 µl, Anachem). The subsequent thermocycling conditions were then applied:

- 98 ºC for 30 seconds
- 10 cycles of:
    - 98 ºC for 10 seconds
    - 60 ºC for 30 seconds
    - 72 ºC for 30 seconds
- 72 ºC for 5 minutes
- Hold at 4 ºC

2.5.12. PCR Clean-Up

AMPure beads were mixed with PCR reaction and washed with ethanol (70%) twice. Then RSB was added (for 5 minutes) to elute DNA from beads. Cleared supernatant (containing final DNA library) was then used in subsequent step.

2.5.13. Library Validation and Preparation

To evaluate library sizes and concentration, the quality control was performed at this step employing Bioanalyzer 1200 assay. After the concentration was calculated, libraries were normalized to 2nM with the Elution buffer (10mM Tris-Cl, pH 8.5 with 0.1% Tween 20). Library pooling for cluster generation and sequencing was then carried out: the equal amount of each normalized sample (2nM) was added to eppendorf tube. To prepare libraries for sequencing DNA was denatured using NaOH (0.2N). HT1 was then added to denatured DNA (8 pM final DNA concentration), and sample was loaded to MiSeq reagent cartridge (and sequenced on MiSeq as above).

2.6. Genome finishing

2.6.1. Direct PCR

Direct PCR was performed using *Elizabethkingia meningoseptica* 501 DNA sample. Oligonucleotides sequences - primers (design method in data analysis chapter; primer sequences in **Appendix I**) were purchased form either Alta biosciences or Sigma-Aldrich. MyTaq DNA polymerase and dNTPs were supplied by Bioline Ltd.

2.6.1.1. Amplification

The primers were diluted to 10 µM with dH_2O. PCR amplification reaction was set in 200 µl PCR tubes. The components were added to PCR reaction tube in the following order (FC – final concentration): nuclease-free H_2O, MyTaq Buffer 5x (1x FC), 25 mM dNTPs (200 µM FC), template DNA (0.08 ng/µl FC), Forward Primer, Reverse Primer (200 nM FC) and 500 units MyTaq DNA Polymerase (0.025 – 0.1 units/µl FC). The reaction was centrifuged and subsequent thermocycling programme was set:

- Initial denaturation: 94 ºC for 30 seconds
- 30 cycles of:
    - Denaturation: 94 ºC for 30 seconds
    - Annealing: 50-58 ºC for 30 seconds
    - Extension: 68 ºC for 1-3 minutes
- Final Extension 68 ºC for 5 minutes
- Hold at 15 ºC.

2.6.1.2. Gel electrophoresis

Buffer:

- ✓ TBE buffer 1x (Tris-Borate-EDTA 10x, pH 8.3, diluted with dH_2O 1:10; AppliChem)

After PCR amplification the presence of DNA products was verified using gel electrophoresis. First, 1.5% gel was made, mixing agarose (Multi-purpose, Bioline) with TBE buffer. Sybr Safe DNA gel strain (1%; Invitrogen) was then added, and the mixture was poured into tank (sealed with tape and with a comb). When the gel was solid TBE buffer was added into tank, and the comb was removed. HyperLadder 1kb (Bioline) was then added to the first and last well. The sample was mixed with DNA loading buffer 5x (Blue, Bioline; 1:5) and added to the sample wells. Electrophoresis was carried out at 100 volts for 60 minutes. PCR products were visualized in transilluminator (Bio-Rad, Molecular Imager Gel Doc XR) employing Quantity One software.

2.6.2. Long Range PCR

To obtain DNA products larger than 6 kb (original Taq PCR limit) Long range PCR was carried out. Accu Taq LA DNA Polymarase kit (Sigma-Aldrich) was used in the experiment. The following reagents were added to the reaction tube (200 µl PCR): nuclease-free H_2O, LA Buffer 10x (1x FC), dNTPs (500 µM FC), template DNA (4 ng/µl FC), DMSO (2% FC), Forward and Reverse Primer (400 nM FC), and Taq LA Polymerase (0.05 units/µl FC). The subsequent thermocycling conditions were used:

- Initial denaturation: 98 ºC for 30 seconds
- 30 cycles of:
    - Denaturation: 94 ºC for 15 seconds
    - Annealing: 55 ºC for 20 seconds

- Extension: 68 ºC for 10 minutes

- Final Extension 68 ºC for 10 minutes
- Hold at 15 ºC.

Gel electrophoresis and visualization was performed as above. PCR products were finally sequenced employing Sanger sequencing.

2.7. Data analysis

The data from Nextera XT and Mate pair sequencing was analysed using Galaxy/Tron software. The reads were aligned using Bowtie 2 alignments to a reference genome. Sequencing alignment data features and properties (such as fragment insert sizes) were visualised employing Qualimap v0.7.1. MIRA (**M**imicking **I**ntelligent **R**ead **A**ssembly) software was used to assemble the reads into contiguous sequences or contigs. MIRA also added Hash Frequency (HAF) tags into the assembly. After the possible joints of the contigs were determined using Gap 5 software and a Contig connection map was created.

Mate pair data was also used to scaffold pre-assembled contigs employing SSPACE software (which assess the order, distance and orientation of contigs). Tablet software was used to determine possible scaffold end connections employing Mate pair information. Primers for PCR genome finishing were designed employing Primer3 (v. 0.4.0) software. Sanger sequencing results were analysed using FinchTV and BLAST (Basic Local Alignment Search Tool). Genome annotation was obtained using RAST (Rapid Annotation using Subsystem Technology). Phylogenetic trees were built in snpTree1. 1 software (the programme is using SNPs – single nucleotide polymorphisms to construct a tree employing whole or assembled genomes) or using BLASTN.

3. Results

3.1. Paired-end whole-genome sequencing (Nextera XT)

Nextera technology is a completely new method designed to prepare nucleic acid samples for sequencing. Nextera XT enables preparation of sequencing-ready libraries for small genomes (bacteria, archaea, viruses), amplicons and plasmids. It is fastest and easiest sample preparation method designed, which requires ultra-low DNA input – only 1 ng. The principle of the Nextera XT involves transposome, which shares and tags (tagment) DNA, adding a special adapter during the tagmentation reaction. Then using this adaptor DNA is amplified employing limited-cycle PCR, which also adds index sequences on both DNA ends, required for cluster generation. Finally, samples are combined and denatured into single strands prior sequencing.

The outcome of previous sequencing experiments (in our lab) was very short, inconsistent length fragments: less than the read length (250 bp). Therefore it was essential to optimize the parameters of DNA preparation for sequencing, such as input DNA concentration (to obtain longer fragments) and incubation time during tagmentation (to ensure that fragments are efficiently tagged and further more effectively amplified).

3.1.1. Quality control (Bioanalyzer)

To prevent sample loss or any sample preparation errors Bioanalyzer quality control was carried out after PCR amplification and PCR clean-up. Agilent Bioanalyzer's a lab on a chip principle employs capillary electrophoresis and utilises fluorescent dye (that binds DNA) to calculate the sizes and concentrations of nucleic acid fragments.

After PCR amplification bioanalyzer analysis showed the large range of fragments with different sizes (150-1500 bp; **Figure 3.1 (sample 1, 2, 3) and Figure 3.2**). The sample with concentration of 0.3 ng showed the best result as the sizes of many fragments were between 500-1500 bp. After PCR clean-up, however, the variety of fragment sizes drastically decreased, leaving only desired long fragments (**Figure 3.1 (sample 7, 8, 9) and Figure 3.3**). At this step all samples showed similar results – one high peak between 600 and 2000 bp.

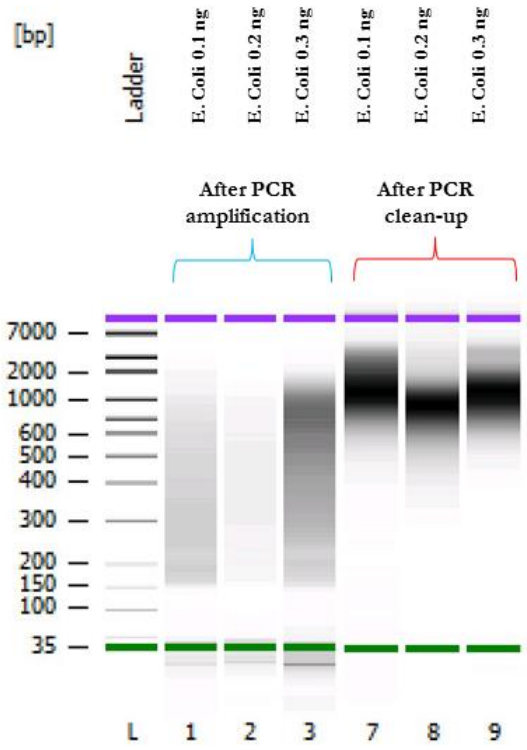


Figure 3.1. Gel image (XT) of *E. coli* samples with different DNA input concentrations (0.1, 0.2, 0.3 ng) after PCR amplification (sample 1-3) and PCR clean-up (sample 7-9). Numbers on the left show DNA insert sizes in base pairs (bp). DNA ladder (L) represents a mixture of DNA fragments of known sizes (based on that samples are aligned to the ladder). Two lines represent lower (▬) and upper (▬) markers – DNA fragments added to each sample, bracketing the DNA sizing analysis. Markers are the internal standards employed to align the ladder with the individual sample analysis.

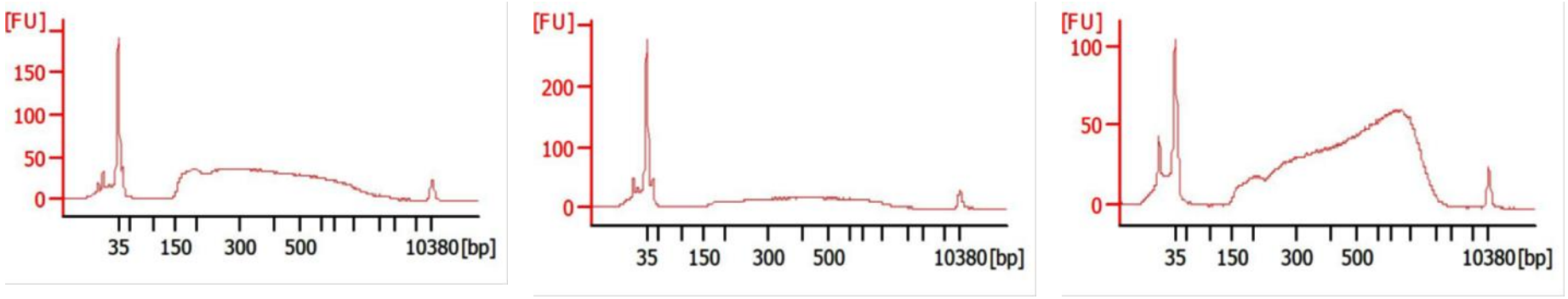


Figure 3.2. Bioanalyzer summary electropherograms (XT1) of *E. coli* samples after PCR amplification. DNA input concentrations 1) - 0.1, 2) - 0.2 and 3) - 0.3 ng. Electropherogram represents Fluorescence intensity (FU) versus Size (bp). Two sharp peaks are lower (35 bp) and upper (10380 bp) markers.

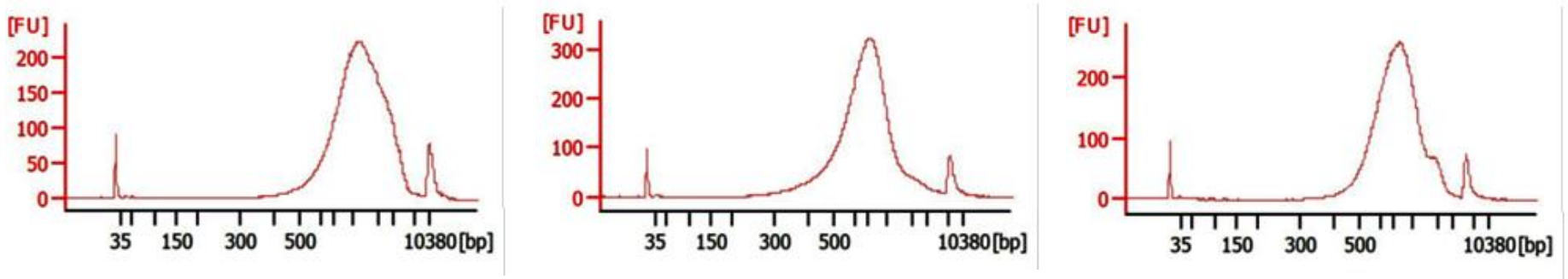


Figure 3.3. Bioanalyzer summary electropherograms (XT2) of *E. coli* samples after PCR clean-up. DNA input concentrations 1) - 0.1, 2) - 0.2 and 3) - 0.3 ng. Electropherogram represents Fluorescence intensity (FU) versus Size (bp). Two sharp peaks are lower (35 bp) and upper (10380 bp) markers.

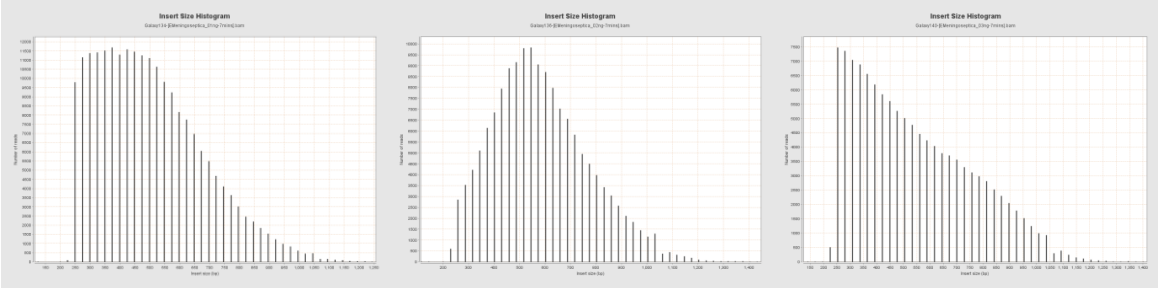
3.1.2. Nextera XT optimization Step 1

During the first optimization stage different sequencing parameters (**Chart 3.1**) – DNA input concentrations (0.1, 0.2 and 0.3 ng/μl) and incubation times (3, 5, and 7 minutes), were applied for 3 bacteria: *Pseudomonas aeruginosa* ST395 specimen 901 (GC content ~65%), *E. coli* O104:H4 specimen 1218_280 (GC ~50%) and *Elizabethkingia meningoseptica* 501 (GC ~38%).

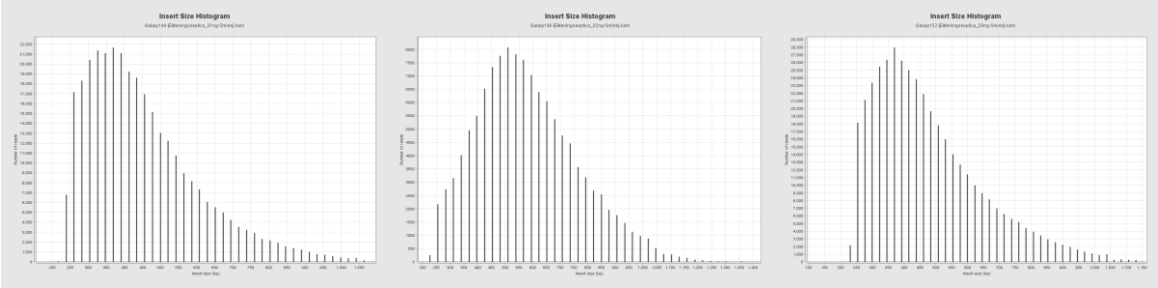
Sample ID	Species	Conditions	Sample ID	Species	Conditions	Sample ID	Species	Conditions
1211	Ec	01ng-7mins	901-1	Pa	01ng-7mins	501-1	Em	01ng-7mins
1218-2	Ec	02ng-7mins	901-2	Pa	02ng-7mins	501-2	Em	02ng-7mins
1218-3	Ec	03ng-7mins	901-3	Pa	03ng-7mins	501-3	Em	03ng-7mins
1218-4	Ec	01ng-5mins	901-4	Pa	01ng-5mins	501-4	Em	01ng-5mins
1218-5	Ec	02ng-5mins	901-5	Pa	02ng-5mins	501-5	Em	02ng-5mins
1218-6	Ec	03ng-5mins	901-6	Pa	03ng-5mins	501-6	Em	03ng-5mins
1218-7	Ec	01ng-3mins	901-7	Pa	01ng-3mins	501-7	Em	01ng-3mins
1218-8	Ec	02ng-3mins	901-8	Pa	02ng-3mins	501-8	Em	02ng-3mins
1218-9	Ec	03ng-3mins	901-9	Pa	03ng-3mins	501-9	Em	03ng-3mins

Chart 3.1. Sample parameters (input DNA concentrations, ng and incubation times, minutes) of the Optimization step 1. Ec – *E. coli*, Pa – *P. aeruginosa*, Em – *E. meningoseptica*.

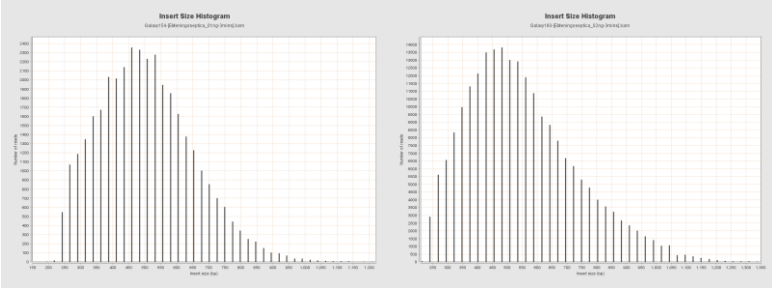
Sequencing results (forward and reverse read for each sample) were imported into Galaxy and aligned using Bowtie 2 to a reference genome (creating a Binary bam alignments file). Insert sizes for each sample were visualised employing Qualimap software. Insert size histograms (**Chart 3.2; Appendix II**) were analysed, and mean, median and mode values were determined for each sample.



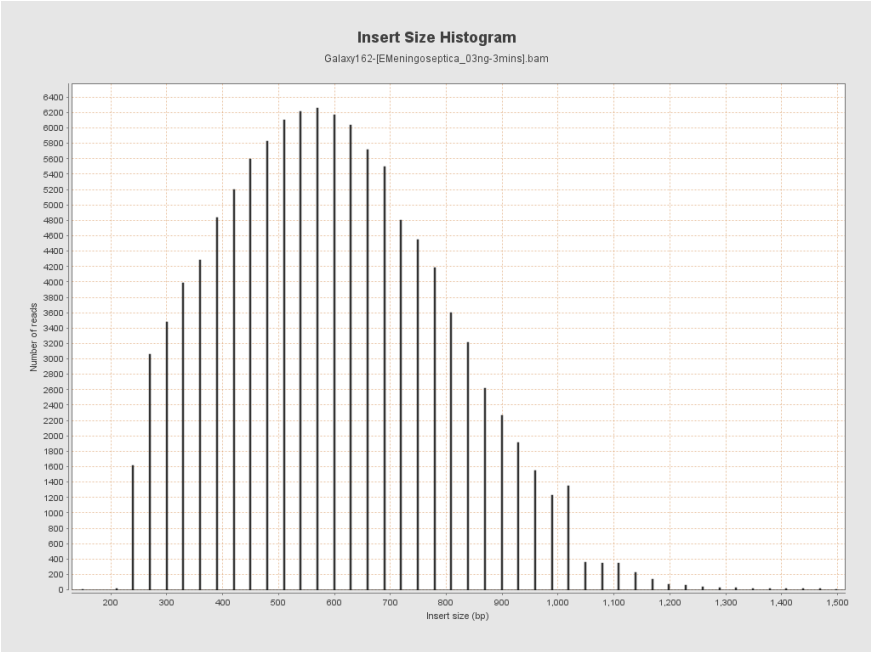
1) 2) 3)



4) 5) 6)



7) 8)



9)

Chart 3.2. Insert size histograms of *E. meningoseptica*. Bars represent number of reads (y axis) at different insert sizes, bp. 1-9 show different sample preparation conditions (listed in Chart 3.1, column 3 in the same order). Created using Qualimap v0.7.1.

The mode insert sizes (averaged) for varying incubation times are showed in **Chart 3.3**. In *E. meningoseptica* samples (a) the average mode value slightly increases when incubation time reduces. In *P. aeruginosa* (b) average mode values remain almost constant. However, in *E. coli* samples (c) mode insert size reaches the highest peak at 5 minutes, but no time-specific trend is observed. Overall, incubation time had no considerable effect on insert sizes during the Optimization step 1. Therefore, 5 minutes incubation time, previously suggested in the Nextera XT protocol, was considered as optimum and used in the subsequent experiments.

Chart 3.4 indicates the mode insert sizes (averaged) for different DNA input concentrations. A similar trend (however not highly significant) is detected in *E. meningoseptica* (a) and *E. coli* (c) mode insert sizes reaching the highest peak at 0.2 ng concentration. The mode insert size in *P. aeruginosa* (b) again does not change significantly (probably due to inaccurate pipetting). To validate these results it was considered to select broader range of concentrations (0.1 – 0.4 ng) and observe the insert sizes in all three bacteria with 5 minutes incubation time (Optimization step 2).

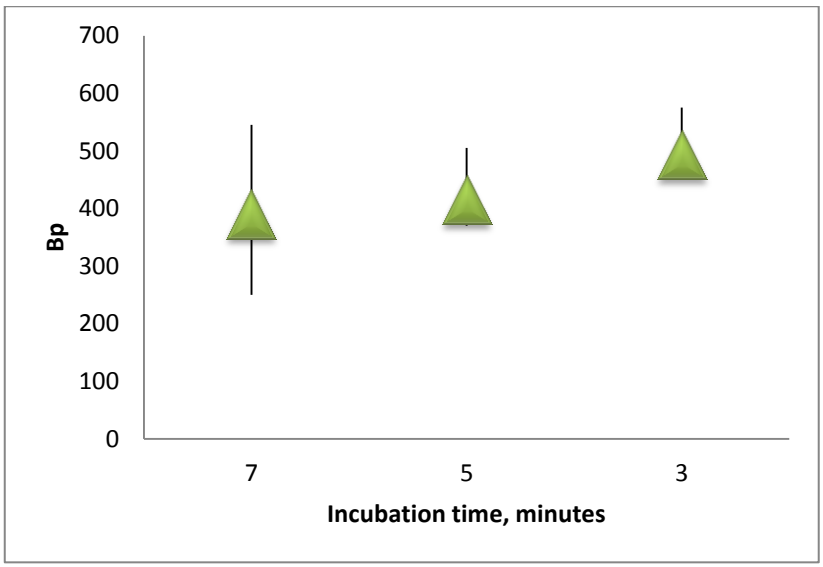
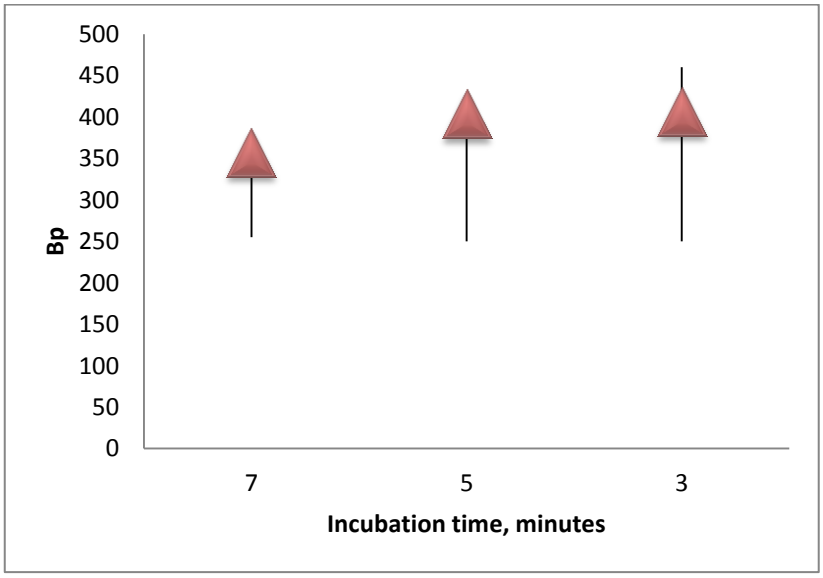
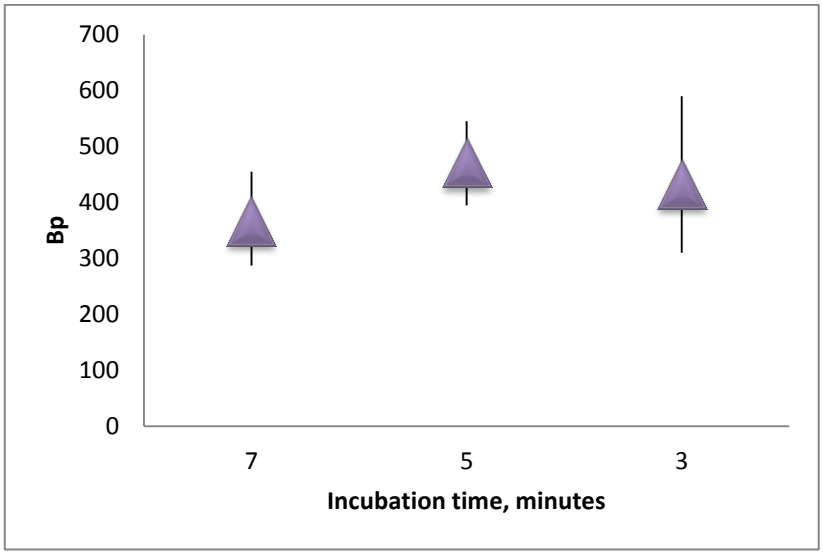


Chart 3.3. Averaged mode insert sizes (base pairs) at different incubation times (3, 5, 7 minutes). Optimization step 1. Data points (Δ) represent average values of 3 measurements.

a) *E. meningoseptica* (Em) mode insert sizes



b) *P. aeruginosa* (Pa) mode insert sizes versus time;



c) *E. coli* (Ec) mode insert sizes versus time.versus time;

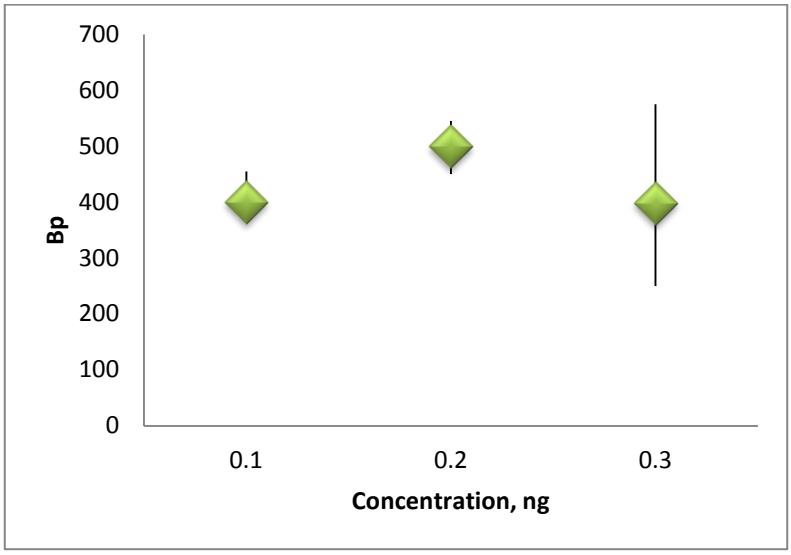
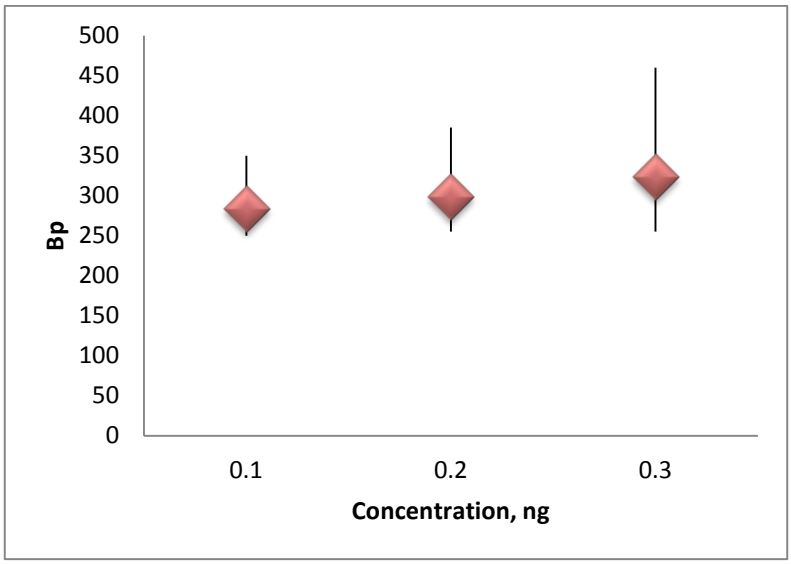
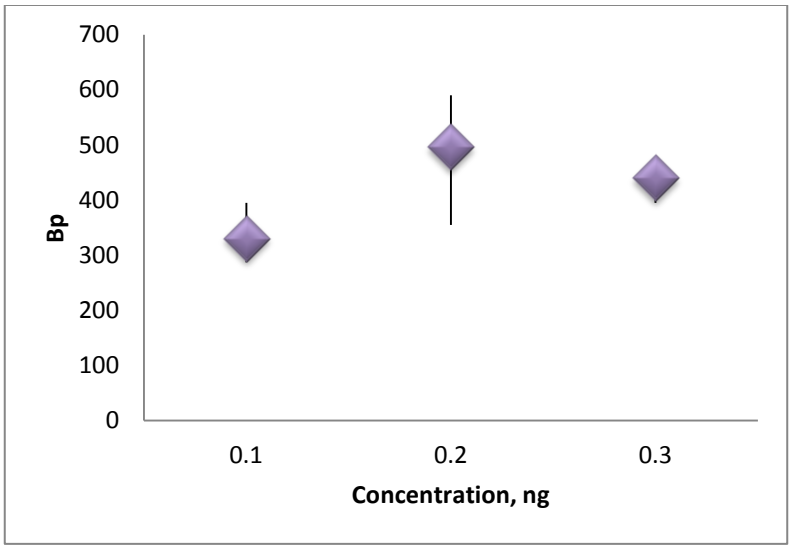


Chart 3.4. Averaged mode insert sizes (base pairs) at different DNA input concentrations (0.1, 0.2, 0.3 ng). Optimization step 1. Data points (◊) represent average values of 3 measurements.

a) *E. meningoseptica* mode insert sizes versus concentration;



b) *P. aeruginosa* mode insert sizes versus concentration;



c) *E. coli* mode insert sizes versus concentration.

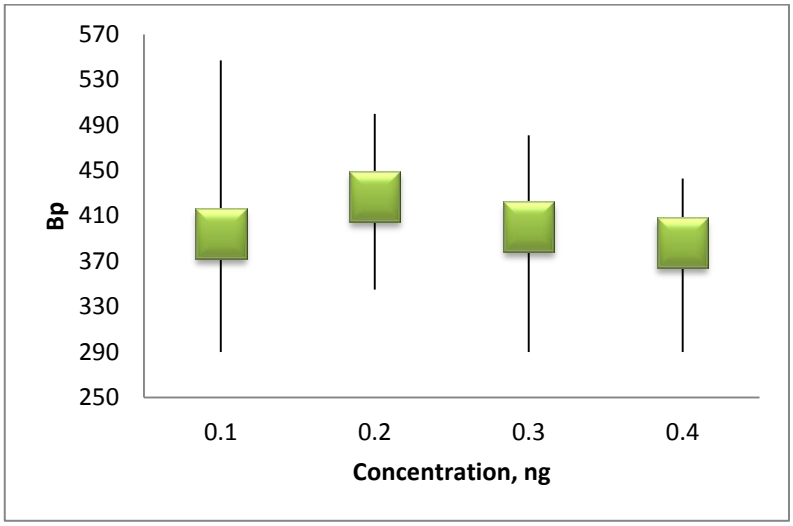
3.1.3. Nextera XT optimization Step 2

To validate the results of Optimization step 1, second Optimization was carried out. At this stage optimum incubation time (5 minutes) and broader range of DNA input concentrations (0.1, 0.2, 0.3 and 0.4 ng/µl) were used with all three bacteria (**Chart 3.5**). This stage also aimed to prevent the possible problems associated with inaccurate pipetting, therefore small amounts of DNA (less than 5 µl) during the dilution step were avoided, as well as only electronic pipettes were used. As in the first stage sequencing results were analysed using Galaxy and visualised employing Qualimap.

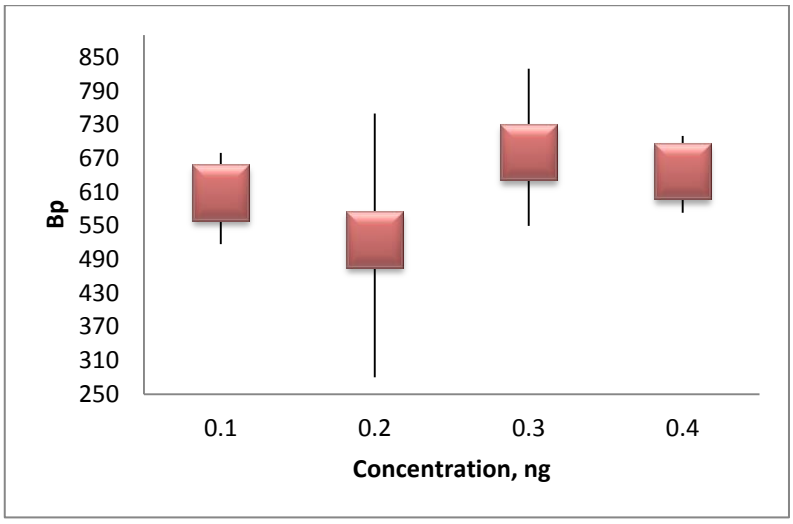
Sample ID	Species	Conditions
1218-10	E. coli	01ng
1218-11	E. coli	02ng
1218-12	E. coli	03ng
1218-13	E. coli	04ng
901-10	P. aeruginosa	01ng
901-11	P. aeruginosa	02ng
901-12	P. aeruginosa	03ng
901-13	P. aeruginosa	04ng
501-10	E. meningoseptica	01ng
501-11	E. meningoseptica	02ng
501-12	E. meningoseptica	03ng
501-13	E. meningoseptica	04ng

Chart 3.5. Sample parameters of the Optimization step 2.

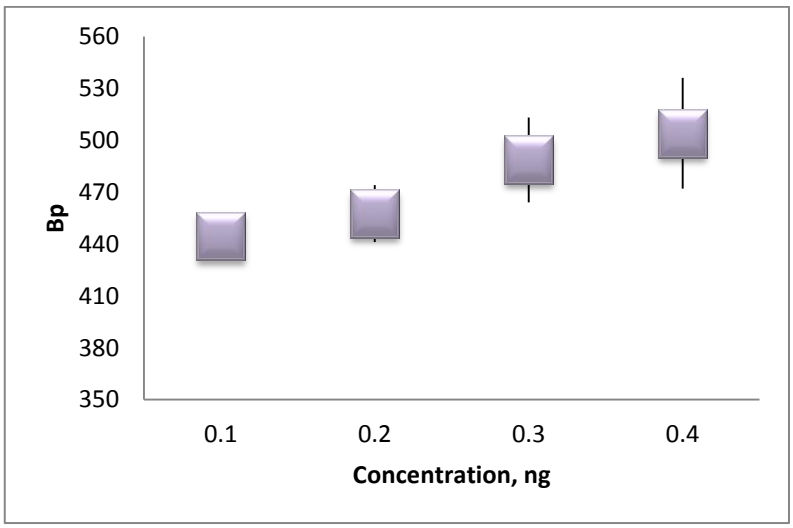
The mode insert sizes (averaged) were calculated form insert size histograms. In *E. meningoseptica* samples (**Chart 3.6,** a) as well as in *P. aeruginosa* (b) no concentration-specific trend is observed. The average mode values randomly fluctuate with the highest peak of 0.2 ng and 0.3 ng for *E. meningoseptica* and *P. aeruginosa* respectively. However, in *E. coli* samples (c) mode insert sizes show significant concentration-dependant trend, considerably increasing along with the concentration. Therefore, DNA input concentration of 0.4 ng was used in subsequent Mate Pair experiments.



a)



b)



c)

Chart 3.6. Averaged mode insert sizes (base pairs) at different DNA input concentrations (0.1, 0.2, 0.3, 0.4 ng) of: a) *E. meningoseptica*, b) *P. aeruginosa* and c) *E. coli*. Optimization step 2**.** Data points (□) represent average values of 3 measurements.

3.1.4. Contig map (1)

After Optimization steps using MIRA software reads were assembled into contigs. The sample with lowest amount of contigs was selected. The best result for *E. meningoseptica* was 0.1ng, 5mins giving 50 contigs in total. Due to the complexity of the genome the best result for *P. aeruginosa* gave over 300 contigs, making manual process of contig map creation almost impossible. Therefore this step was performed only for *E. meningoseptica* (*P. aeruginosa* contig amount were further reduced in Mate Pair sequencing). To start the process of contig joining (as well as the process of genome finishing) each end of the contig (mostly containing multiple linkage options) was examined using Gap 5 software. To evaluate the coverage of each end the software used HAsh Frequency (HAF) tags, which display the status of read parts in comparison to the whole genome (**Figure 3.4**). The following HAF tags are applied to the fragments based on coverage: HAF2 – coverage below average, HAF3 – coverage is at average, HAF4 – coverage above average, HAF5 – probably repeat, HAF6 – 'heavy' repeat, HAF7 – 'crazy' repeat. Consequently based on Gap 5 results the initial Contig map (**Figure 3.5**) was created indicating all possible connections between contigs.

Figure 3.4. Gap 5 screenshot - part of the contig 1. Colours represent fragments tagged with different HAF tags. HAF3 – green, HAF4 – yellow, HAF5 – red.

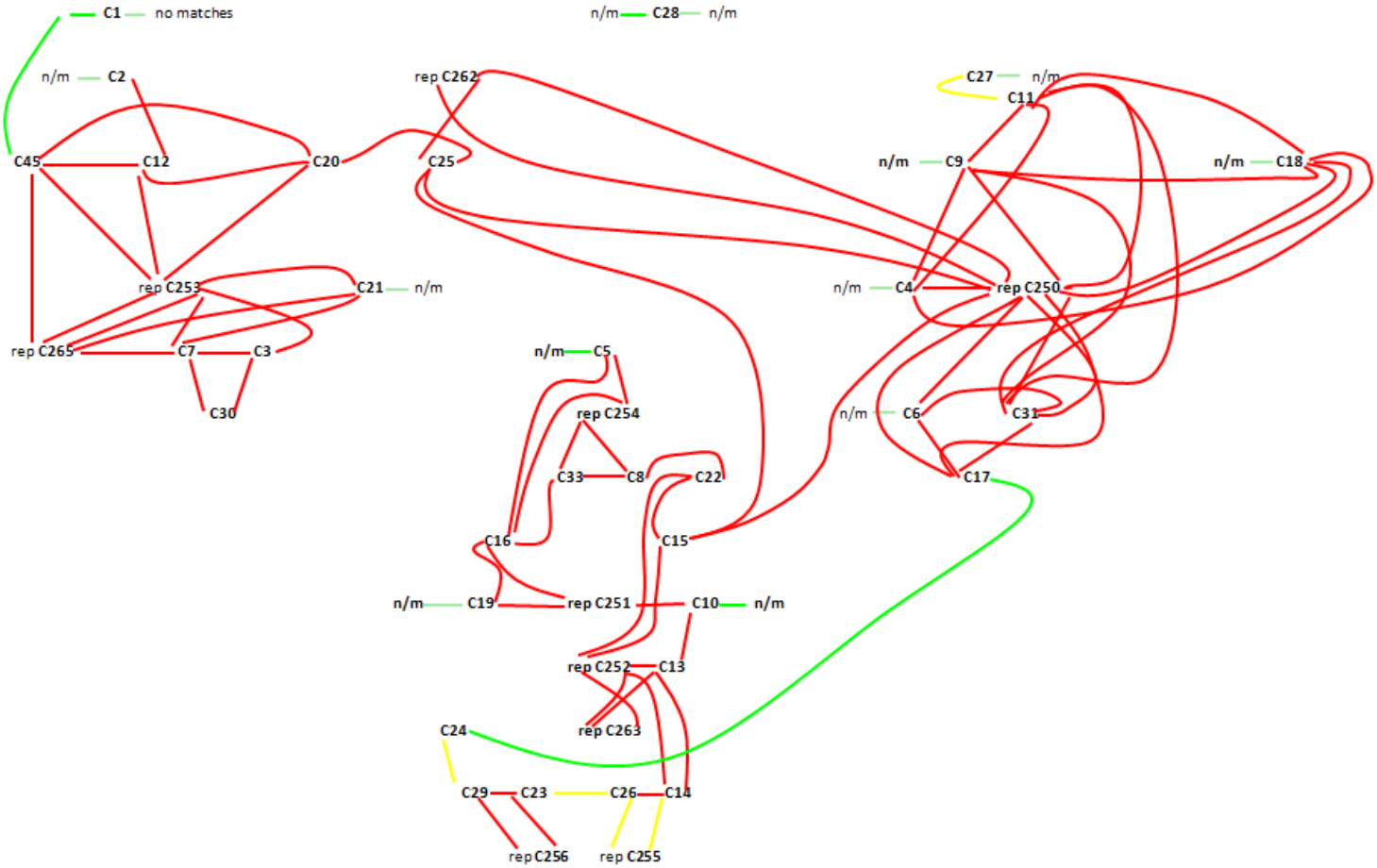


Figure 3.5. Nextera XT result based contig map (Created in MS Office Word, using Gap 5 data). C – contig, rep – repeat, n/m – no matches found in Gap 5. Coloured lines represent connections between contigs, marked with HAF tags. HAF2 – light green, HAF3 – green, HAF4 – yellow, HAF5 – red.

3.2. Mate pair sequencing

To obtain more precise information about contigs as well as to join them into scaffolds Nextera Mate Pair sequencing was carried out. Mate Pair sample preparation technology allows the creation of libraries with long insert sizes ranging from 2 to 12 kb. Using Mate Pair sequencing in addition to conventional paired-end sequencing creates a very powerful strategy for many applications including genome finishing.

Nextera Mate pair principle is similar to Nextera XT, however is more complex and has additional features that enhance sequencing results. It involves specific mate pair transposome, which tagment DNA, adding a biotin junction adapter. This adapter is later used to detect the location where the ends of fragments have joined to form a mate pair. DNA is then circularized in blunt ended intramolecular ligation. Circularized molecules are physically shared, and further biotin-selected. TruSeq indexing adapters are then added to the ends of mate pair fragments, and are used in following DNA amplification.

3.2.1. Quality control (Bioanalyzer)

Quality control of mate pair fragments was performed after AMPure Purification of Strand Displacement reaction and after PCR clean-up. After Purification of Strand Displacement reaction all samples (**Figure 3.6,** 1-4; **Figure 3.7**) showed same results – 2000 to 10 000 bp long, good purity and high concentration fragments.

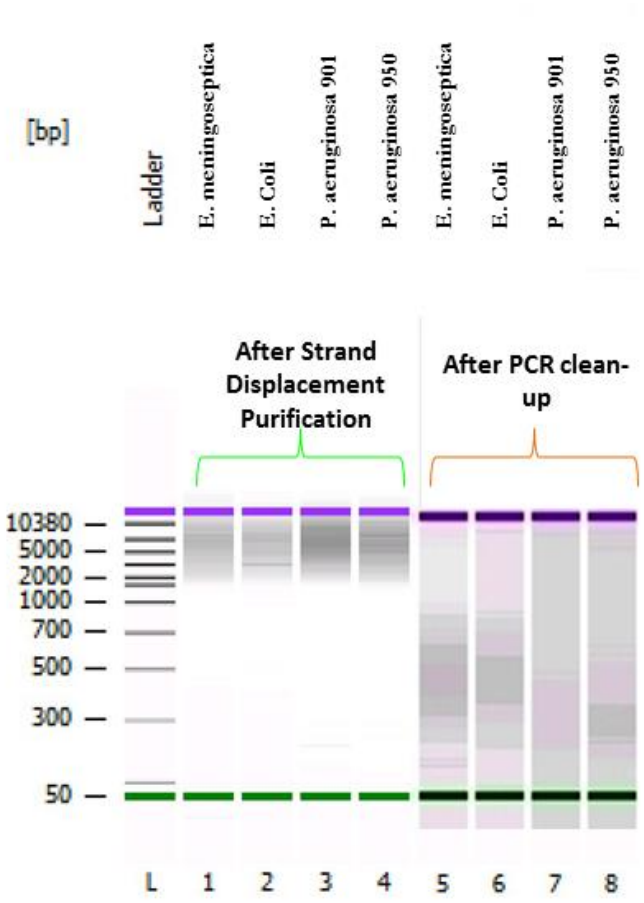


Figure 3.6. Gel image (MP) of *Elizabethkingia meningoseptica, E. coli*, *Pseudomonas aeruginosa* 901 and 950 after AMPure Purification of Strand Displacement reaction (sample 1-4) and after PCR clean-up (sample 5-8). Numbers on the left show DNA insert sizes in base pairs (bp). L – DNA ladder. Two lines represent lower (▬) and upper (▬) markers.

After PCR clean-up Bioanalyzer analysis showed that fragment sizes in different samples varied (ranging from 240 to 700 bp; **Figure 3.6,** 5-8; **Figure 3.8**). According to Mate pair protocol final size of libraries, obtained using Gel-free method, should be in a range between 300 and 1500 bp. The molarity of the sample should be between 5 and 50 nM. **Chart 3.7** indicates that *E. meningoseptica* showed the best Mate pair sample preparation results with mode library size of 500 bp. Both *P. aeruginosa* 901 and P. *aeruginosa* 950, however, (most likely) experienced some sample loss, therefore library sizes were slightly lower than optimum value. Nevertheless, the molarity of all samples was optimal.

No	Sample	Peak size, bp	Conc. [ng/µl]	Molarity, [nmol/l]
5	*E. meningoseptica*	503	6.44	19.4
6	*E. coli*	385	5.22	20.6
7	*P. aeruginosa* 901	278	2.58	14.1
8	*P. aeruginosa* 950	240	3.71	23.4

Chart 3.7. Bioanalyzer peak (mode insert size) data. Mate pairs PCR cleaned-up samples. Conc – Concentration.

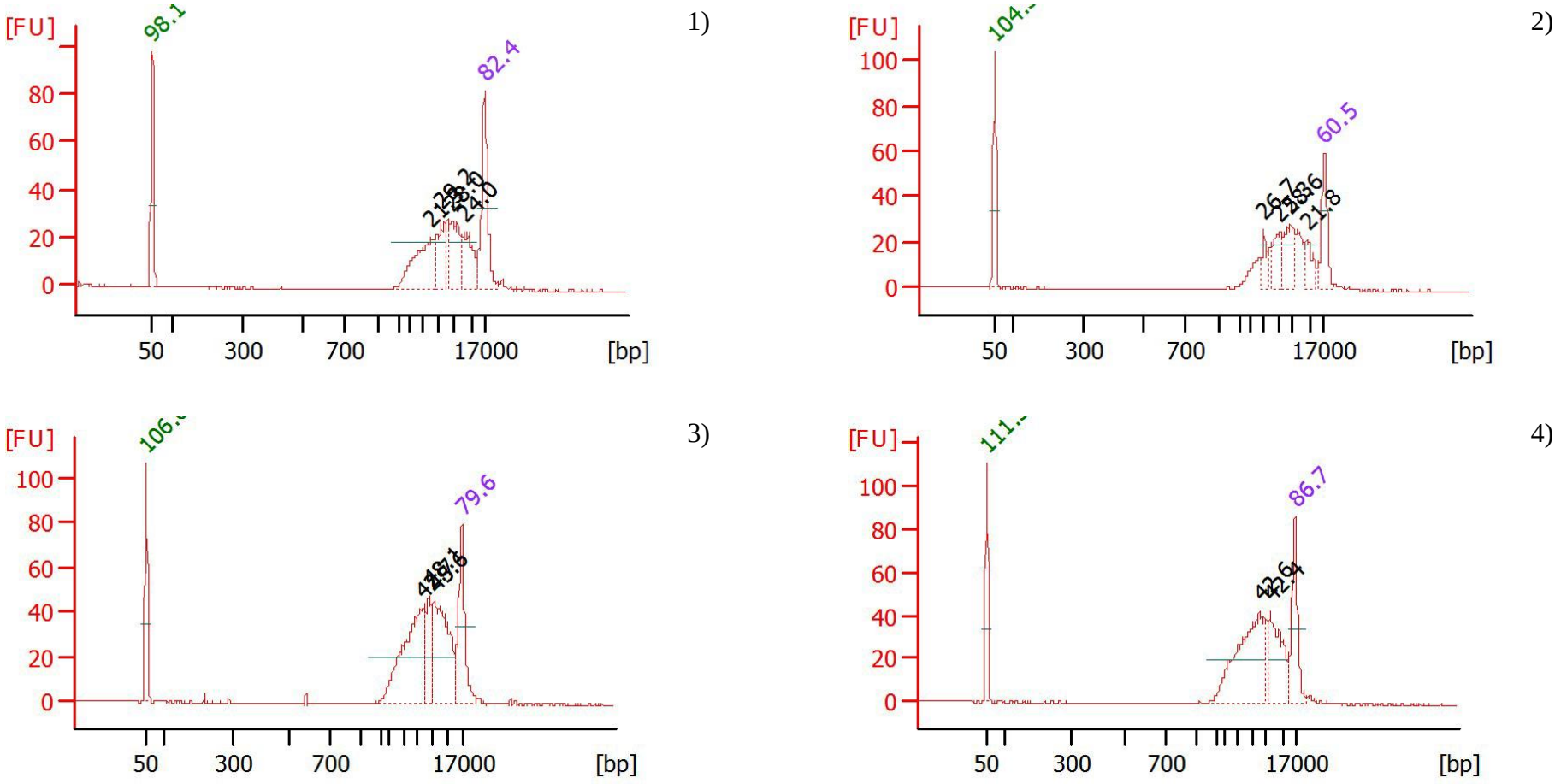


Figure 3.7. Bioanalyzer summary electropherograms (MP1) of: 1) *E. meningoseptica*, 2) *E. coli*, 3) *P. aeruginosa* 901 and 4) *P. aeruginosa* 950 after AMPure Purification of Strand Displacement reaction. Electropherogram represents Fluorescence intensity (FU) versus Size (bp). Two sharp peaks are lower (35 bp) and upper (10380 bp) markers.

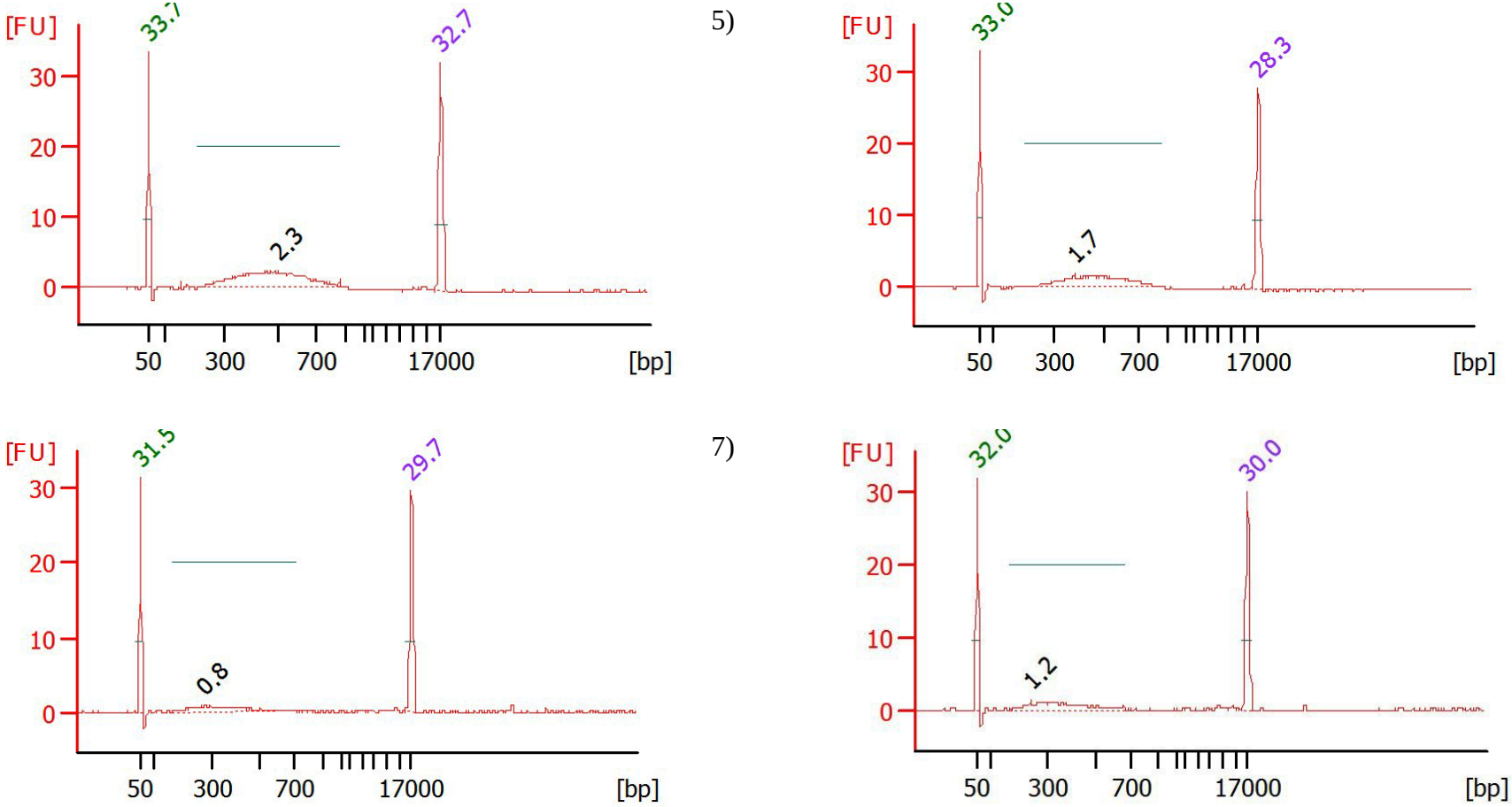
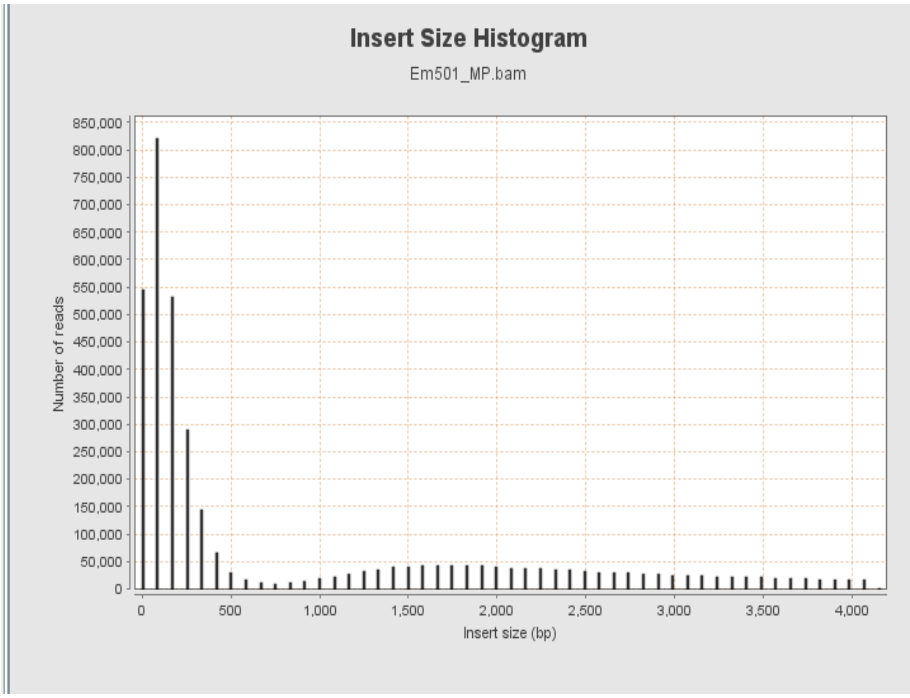


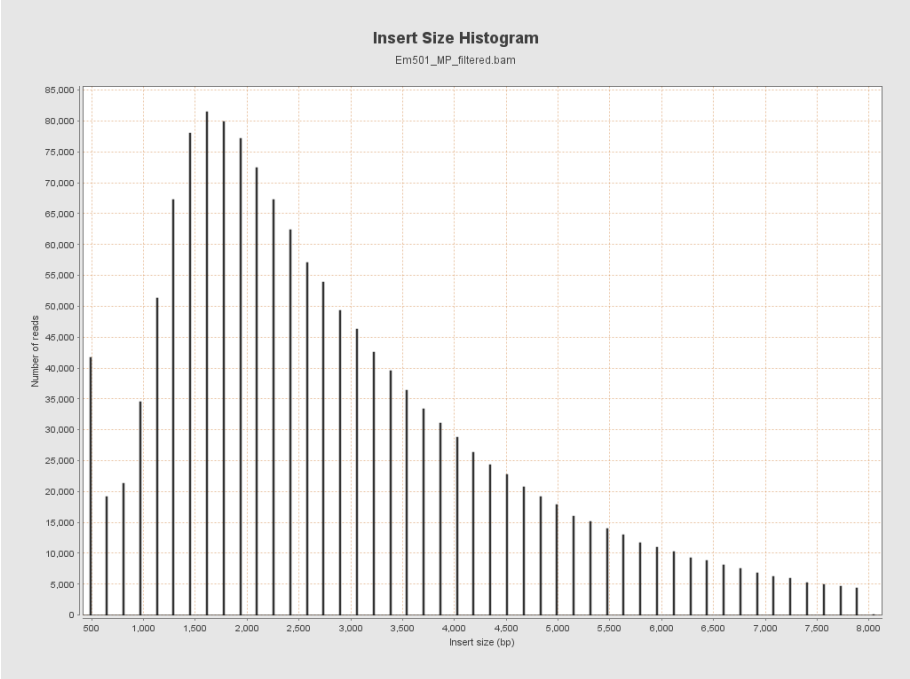
Figure 3.8. Bioanalyzer summary electropherograms (MP2) of: 5) *E. meningoseptica*, 6) *E. coli*, 7) *P. aeruginosa* 901 and 8) *P. aeruginosa* 950 after PCR clean-up. Electropherogram represents Fluorescence intensity (FU) versus Size (bp). Two sharp peaks are lower (35 bp) and upper (10380 bp) markers.

3.2.2. Mate pair results

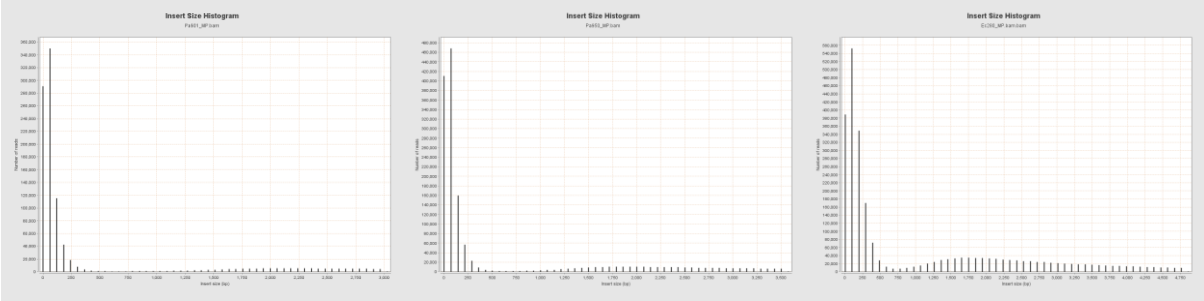
The results of Mate pair sequencing were large amount of short paired-end fragments and wide range of long mate-paired fragments (**Chart 3.8)**. Mate Pair sample preparation kit considerably enhanced sequencing results increasing the mode insert sizes from average 450 bp to 1800 bp. The mode insert size (**Chart 3.9 a**) for *E. meningoseptica*, was approximately 1600 bp, for *E. coli* – 1700 bp and for *P. aeruginosa* – 2100 (901) and 1900 bp (950). In Nextera XT 2nd optimization step as well as in Mate Pairs insert sizes increased along with GC content. The average coverage (**Chart 3.9 b**) also considerably increased (5-7 times) compared to paired-end results. However, both paired-end and Mate Pair results indicated a similar trend – coverage was the highest in *E. meningoseptica* and lowest in *P. aeruginosa* (and *E. coli* in the middle). This trend was most likely observed due to the size of the genomes (*P. aeruginosa* has the largest genome and *E. meningoseptica* – smallest) or due to different GC content (highest in *P. aeruginosa* and lowest in *E. meningoseptica*), or both. Moreover, Mate Pair sequencing considerably simplified scaffolding and reduced finishing effort for *E. meningoseptica* genome.



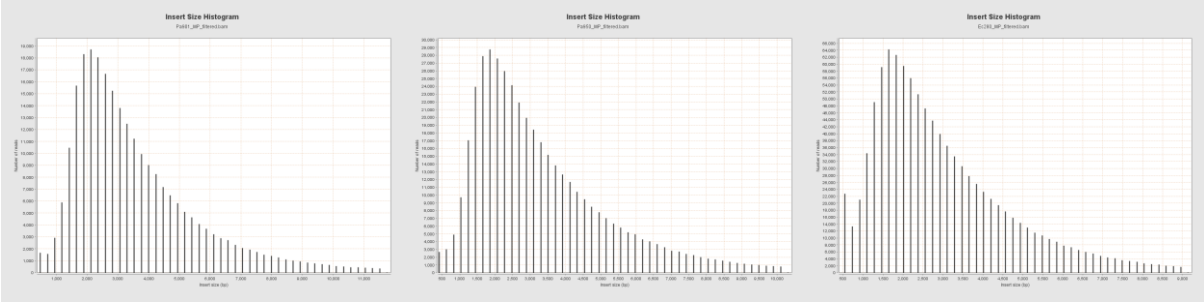
a)



b)

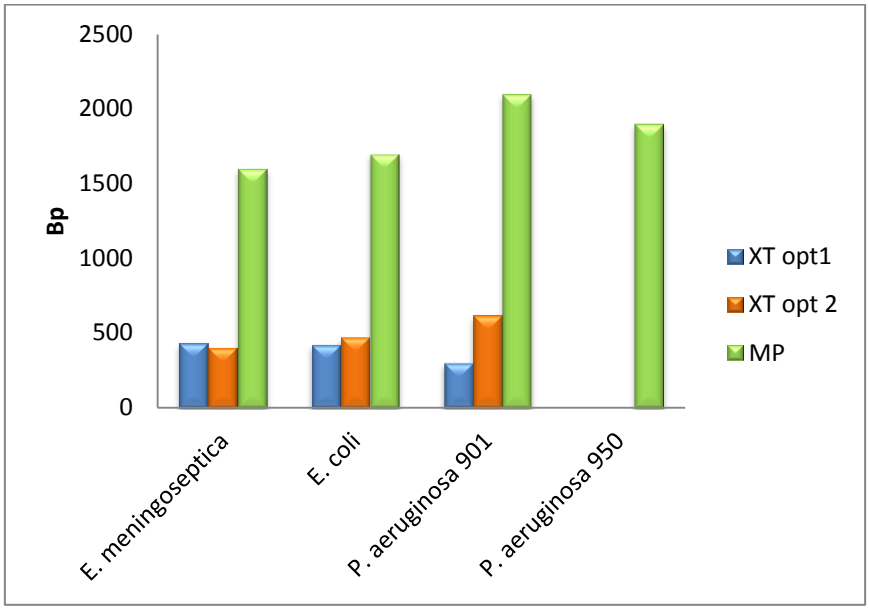


c) d) e)

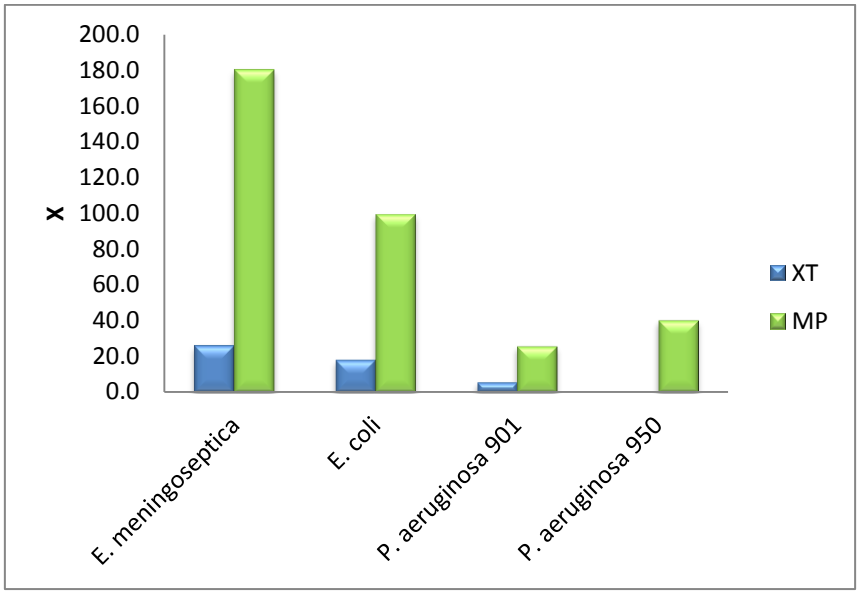


f) g) h)

Chart 3.8. Mate Pair insert size histograms. Histograms a (*E. meningoseptica*), c (*P. aeruginosa* 901), d (*P. aeruginosa* 950), and e (*E. coli*) represent overall Mate Pair results, where high peaks between 0 and 500 bp represent paired-end fragments, lower peaks from 500 bp – mate pairs. Histograms b (*E. meningoseptica*), f (*P. aeruginosa* 901), g (*P. aeruginosa* 950), and h (*E. coli*) show filtered mate pair fragments across different insert sizes.



a)



b)

Chart 3.9. Mode insert sizes (base pairs; **a**) and average coverage (x – times; **b**) for *E. meningoseptica*, *E. coli* and *P. aeruginosa* (901 and 950) after Nextera XT first optimization step (blue), second optimization step (orange) and Mate Pair sequencing (green).

3.2.3. Scaffolding

During this step scaffolding was performed using SSPACE software, which joined 50 contigs of *E. meningoseptica* into 24 scaffolds (number of scaffolds for *P. aeruginosa* still remained high (over 200), requiring further optimization and computational processing). Main scaffolds No 1-10 and 13 (**Chart 3.10**) were non-repetitive sequences (excluding some repeats already

assembled in scaffolds), which together form the chromosome with the gaps in between. Repetitive fragments (No 11, 12, 14-24; **Chart 3.11**) subsequently should fill in the gaps between non-repetitive contigs, thereby making the whole genome.

Scaffold	Contigs in the scaffold	Gap size, bp	Scaffold	Contigs in the scaffold	Gap size, bp
Scaffold 1	c5	gaps-472	**Scaffold 5**	c17	gaps-1367
	c2	gaps-400		c24	gaps-458
	c12	gaps177		c29	gaps-632
	rep c265	gaps-1198		rep c256	gaps-320
	c21	gaps-376		c23	gaps-493
	c28	gaps-124		c26	gaps-546
	c4	gaps-693		c14	
	rep c250	gaps-899	**Scaffold 6**	c45	gaps-51
	c25	gaps-551		c1	gaps-813
	c20	gaps-719		c10	gaps87
	c7			rep c251	gaps-298
Scaffold 2	c19	gaps-692		c16	gaps-206
	c27	gaps-592		c33	gaps164
	c11			c18	
Scaffold 3	c3		**Scaffold 7**	c13	
Scaffold 4	c8	gaps-481	**Scaffold 8**	c15	
	c22		**Scaffold 9**	c9	gaps-289
				c6	
			Scaffold 10	c30	
			Scaffold 13	c31	

Chart 3.10. Main scaffolds created with SSPACE. C – contig.

Repetitive fragments	Contigs	Repetitive fragments	Contigs
Scaffold 11	rep c252	Scaffold 19	rep c259
Scaffold 12	rep c253	Scaffold 20	rep c258
Scaffold 14	rep c254	Scaffold 21	rep c257
Scaffold 15	rep c262	Scaffold 22	rep c260
Scaffold 16	rep c263	Scaffold 23	rep c266
Scaffold 17	rep c255	Scaffold 24	rep c261
Scaffold 18	rep c264		

Chart 3.11. Repetitive fragments. Scaffolds created with SSPACE. Rep – repeat.

Scaffolding process allowed facilitating the initial Contig map creating a Simplified map (No 2; **Figure 3.9**). Most of contigs were now certainly connected to each other making scaffolds (leaving only a few contigs without scaffolded connections), what considerably reduced possible variations, however, the scaffold ends still had a number of probable connections. In the next steps it was attempted to join these scaffold ends together and fill in the gaps in between contigs.

Figure 3.9. Simplified Contig map (No2). Green lines – scaffolded connections, red and yellow – possible connections (Gap5 determined), light green lines – no matches (N/M) determined. C – contig.

3.3. *E. meningoseptica* Genome finishing

3.3.1. Direct PCR. Preliminary data.

In order to finish the genome of *Elizabethkingia meningoseptica* 501 – fill in the gaps between scaffolded contigs generated during Mate pair sequencing as well as to estimate non-scaffolded connections (and to join the scaffolds), direct PCR was performed. Reactions estimating the connection between 2 contigs (or validating the connection and filling in the gap if contigs are in the scaffold) were created based on the second (simplified) Contig Map. Firstly, PCR was carried out using the protocol suggested conditions (0.025 units/µl of Taq Polymerase, 58 ºC annealing temperature, and 1 minute extension time). During this stage only one connection was detected (Reaction No 27; Reaction list in **Appendix III)** employing Gel Electrophoresis. However, this reaction Electropherogram showed very weak band (Figure not shown), therefore PCR Optimization was carried out.

3.3.2. PCR Optimization

PCR Optimization was performed for the Reaction No 27 (**Figure 3.10**), varying the amount of Taq Polymerase (0.025 – 0.1 units/µl), lowering the annealing temperature (which was estimated to be too high for primers) from 50-56 ºC and changing the extension time to 3 minutes (to give enough time for building required product). Sample number **5** (0.05 units of Taq, 52 ºC annealing temperature) showed the best result during Optimization step, therefore this conditions were used in subsequent PCR experiments.

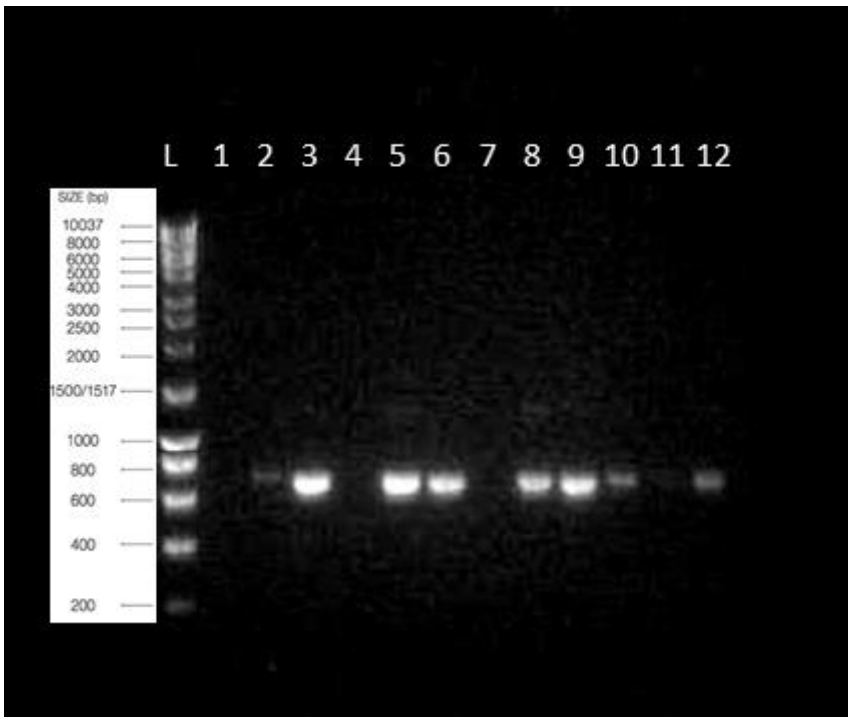


Figure 3.10. PCR Optimization Gel image. Reaction No 27. L – DNA Ladder, 1-12 – samples. Sample conditions (First number – Polymerase in units/µl, Second number – annealing temperature in ºC): **1** – 0.025, 50; **2** – 0.05, 50; **3** – 0.1, 50; **4** – 0.025, 52; **5** – 0.05, 52; **6** – 0.1, 52; **7** – 0.025, 54; **8** – 0.05, 54; **9** – 0.1, 54; **10** – 0.025, 56; **11** – 0.05, 56; **12** – 0.1, 56. Fragment sizes (according to ladder) represented in base pairs (bp). Agarose gel 1.5%, stained with Sybr Safe.

3.3.3. PCR

At this stage all PCR reactions were carried out using 0.05 units/µl of Taq Polymerase, 52 ºC annealing temperature and 3 minutes extension time. Almost all connections between contigs in scaffolds (Reaction No 1-5, 7-11, 13-17, 19-22 and 40; excluding reactions, where product was larger than ~ 6 kb – Taq Polymerase limit; for these reactions Long range polymerase was used) were validated after optimization (**Figure 3.11**). Moreover, new connections were discovered (Reaction No 24, 27, 28, 47, 56, 75, 80), linking scaffolds together. Furthermore, to obtain sequences of the products (to fill in the gaps and validate new connections) Sanger sequencing was performed.

Figure 3.11. PCR (optimized) Gel image. Reaction No 1-17. L – DNA Ladder, 1-17 – samples. Fragment sizes (according to ladder) represented in base pairs (bp). Agarose gel 1.5%, stained with Sybr Safe.

3.3.4. Long-range PCR

To obtain the products larger than 6 kb Long range (LR) PCR was carried out. Conditions used in LR PCR were protocol suggested – Taq LA Polymerase, 0.05 units/µl final concentration and thermocycling conditions mentioned in methods. However, the products failed to develop (even a positive control) as the conditions of LR PCR were unsuitable (therefore, LR PCR requires future optimization).

3.3.5. Third Contig map

After the PCR step scaffolded connections were validated and 7 new connections were obtained. That enabled the creation of the Third Contig map (**Figure 3.12**), and left only 4 connections undetermined, giving 7 possible variants (determined employing Tablet). All

contigs were then joined into appropriate scaffolds (according to **Chart 3.10**) giving the final scaffold map (**Figure 3.13**).

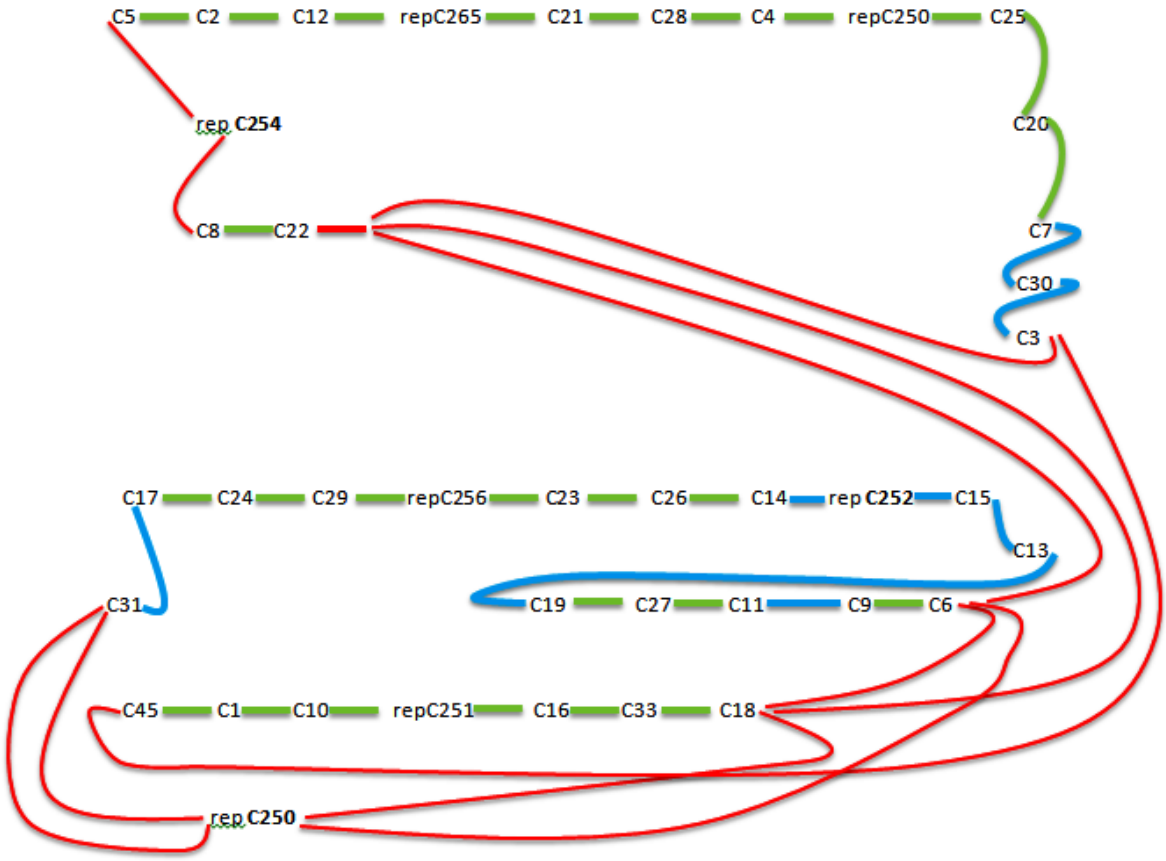


Figure 3.12. Third Contig map. Green lines – scaffolded connections, red – possible connections (Tablet determined), blue – PCR determined connections. C – contig.

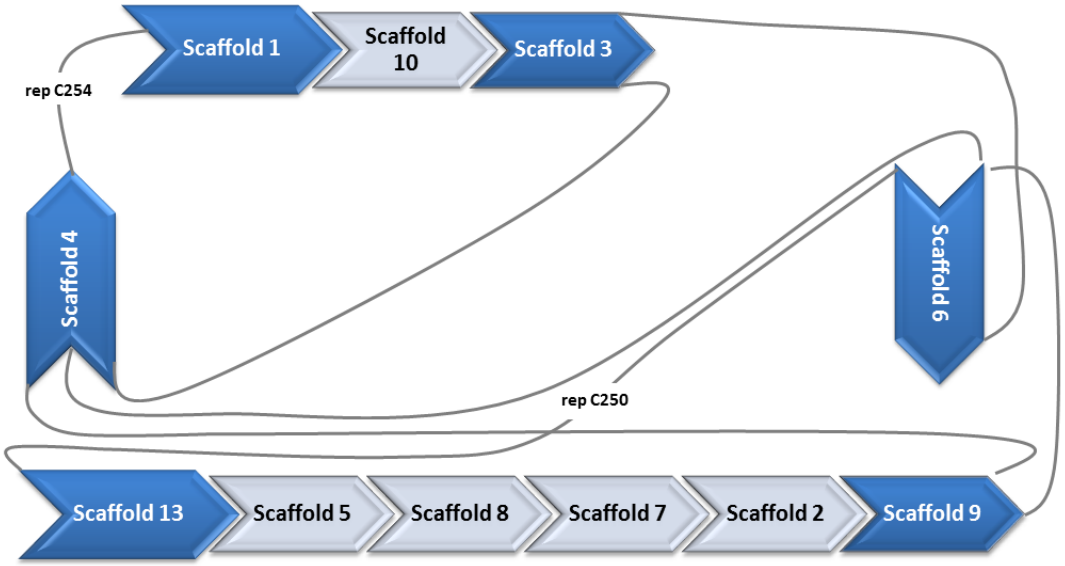


Figure 3.13. Scaffold map. Blue figures represent scaffolds with free ends, grey – with joined ends. Lines – still possible connections (according to Tablet). Rep – repetitive contig.

3.3.6. Sanger sequencing

Sanger sequencing was the final stage of genome finishing. It was used to validate PCR products by determining their sequence and aligning the sequence using BLASTN (in Galaxy software) to scaffoldded sequences of *E. meningoseptica*. Successful Sanger sequencing results (clean from noise sequences) were obtained with almost all samples (5 reaction products were sequenced – No 17, 47, 56, 75, 80, each with 2 primers from either 3' (a) or 5' end (b) giving 10 samples (1 - 17a, 2 - 17b, 3 - 47a and so on) in total), excluding reaction No 47 (a and b), where high frequency noise was detected. In all other samples the identity ratio with the estimated contigs was from 97-100%, what validated these connections between contigs.

3.4. Annotation

3.4.1. Genome features

To obtain genome annotations assembled bacterial sequences were analysed in RAST software. Genome size of *Elizabethkingia meningoseptica* was evaluated for the first time in this study and was estimated to be 4,173,270 bp. Genome sizes of *Pseudomonas aeruginosa* varied drastically (however, still were in predicted range) being 5,943,010 bp for specimen 901 and 6,545,995 bp for 950. *E. coli* genome size was 5,407,699 bp. GC content was estimated to be 36, 66, 66 and 51% for *E. meningoseptica*, *P. aeruginosa* 901, *P. aeruginosa* 950 and *E. coli*, respectively. Number of coding sequences for *E. meningoseptica*, *P. aeruginosa* 901, *P. aeruginosa* 950 and *E. coli* were 3858, 5444, 6066 and 5492; and number of RNAs 54, 64, 61 and 128, respectively.

3.4.2. Subsystems

RAST software divided all genomes into 27 subsystems (**Chart 3.12-3.15**), which were further allocated in 15 categories: Cofactors, Vitamins, Prosthetic Groups, Pigments; Membrane and membrane transport; Resistance; Metabolism; Mobile elements; Nucleosides and Nucleotides; Cell cycle; Regulation; Lipids; Respiration; Stress Response; Amino Acids and Derivatives; Carbohydrates; Motility and Chemotaxis; Miscellaneous.

In all 4 bacteria analysed Metabolism was the largest category comprising 25-27% of the genome. Percentage ratio of two other large categories – Amino Acids and Derivatives and Carbohydrates was similar in *Elizabethkingia meningoseptica* (18% and 14%), *Pseudomonas aeruginosa* 901 (20% and 9%) and 950 (16% and 9%), however, *E. coli* showed completely opposite trend (9% and 18%). Cofactors, Vitamins, Prosthetic Groups, Pigments were represented by 6-9%. The most important category in pathogenic bacteria – Resistance

comprised 3-5% of the genome. Another groups also important in pathogens – Membrane and membrane transport, and Stress Response was represented by 8-11% and 3-5% respectively. Motility and Chemotaxis group was present in all bacteria, excluding *E. meningoseptica* as it is non-motile organism. Other categories were quite small – represented in all bacteria by 5% or less.

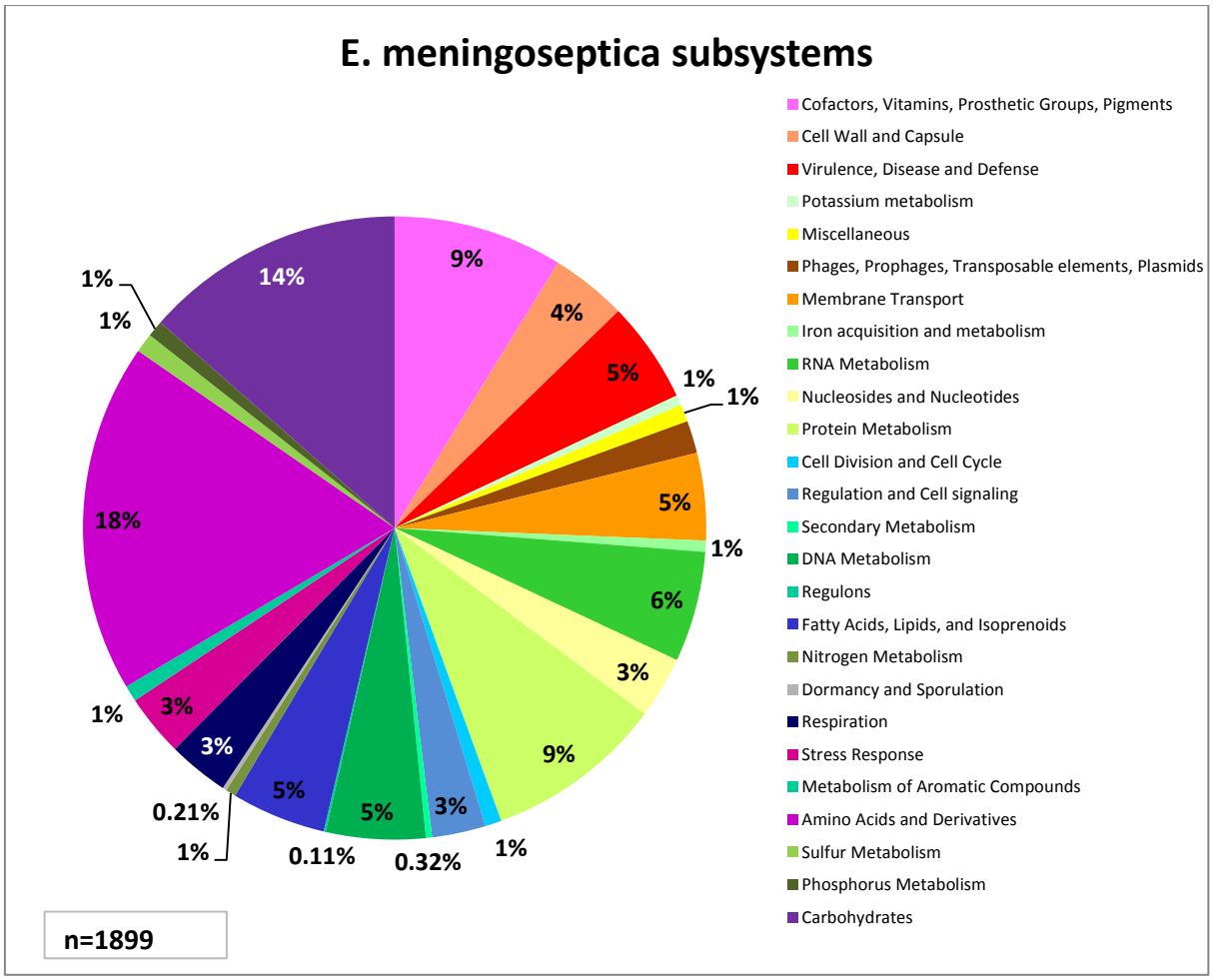


Chart 3.12. *E. meningoseptica* subsystems. Each subsystem is represented with different colour and its percentage ratio to the whole genome is showed. N – total number of features in subsystems.

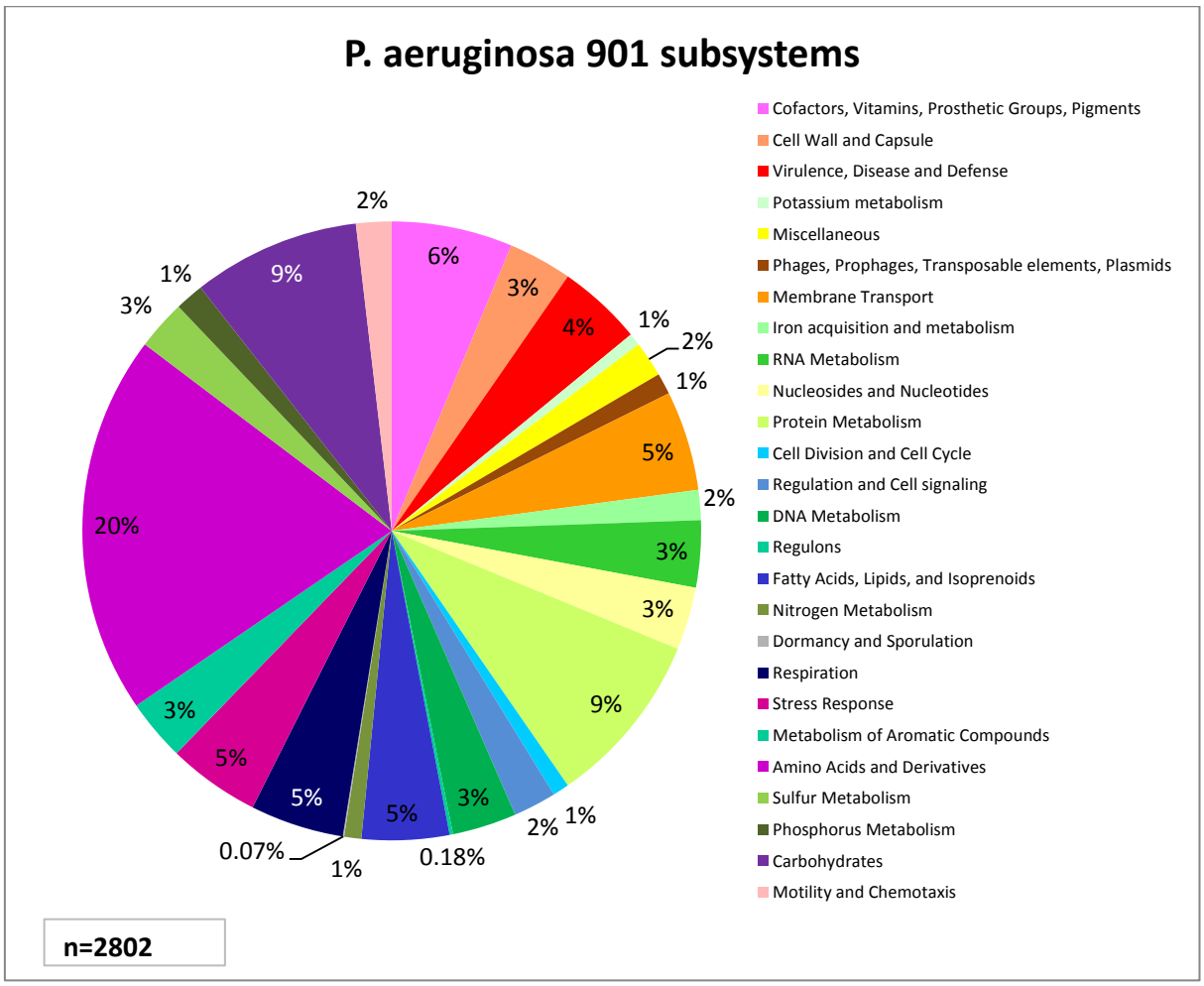


Chart 3.13. *P. aeruginosa* 901 subsystems. Each subsystem is represented with different colour and its percentage ratio to the whole genome is showed. N – total number of features in subsystems.

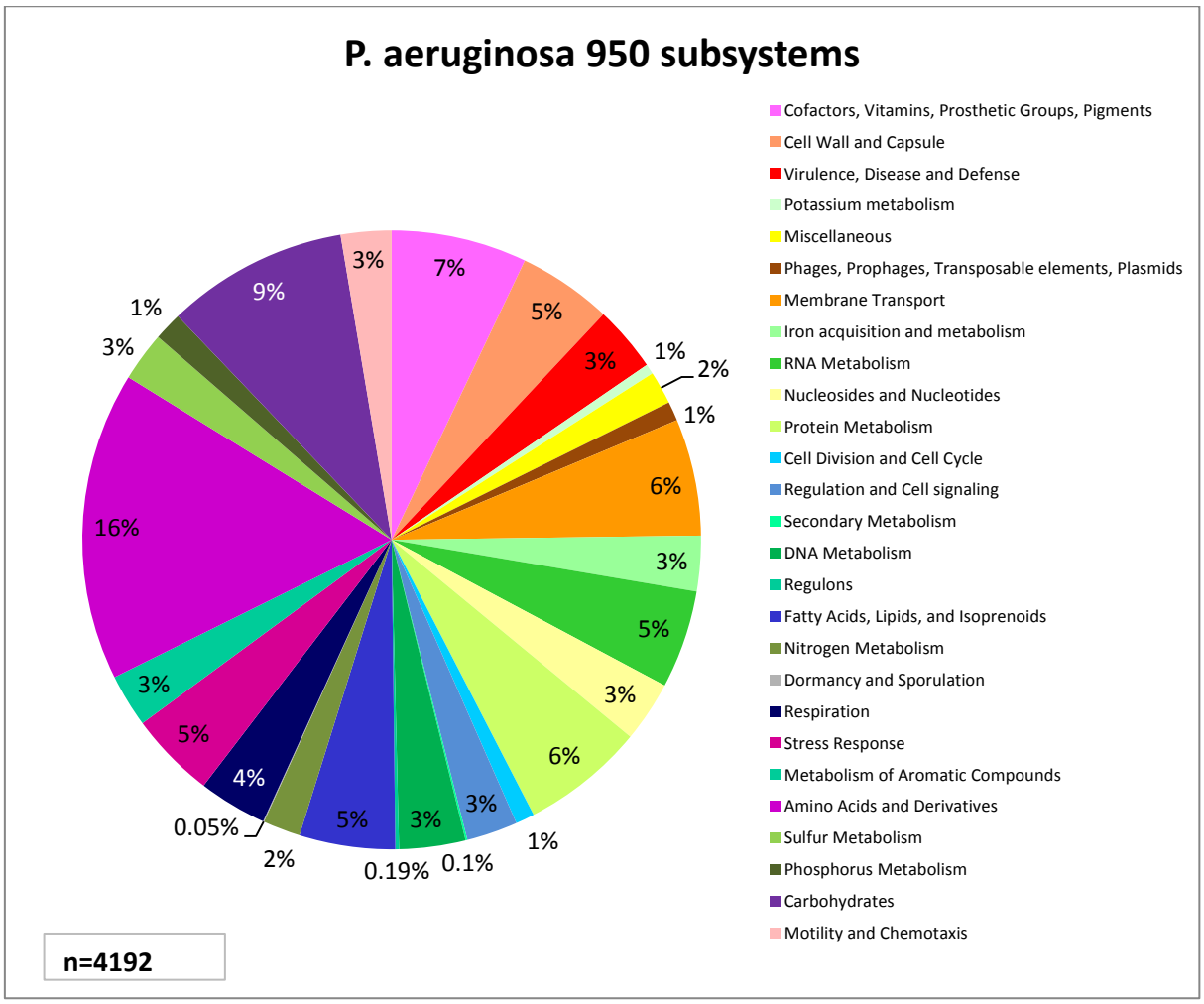


Chart 3.14. *P. aeruginosa* 950 subsystems. Each subsystem is represented with different colour and its percentage ratio to the whole genome is showed. N – total number of features in subsystems.

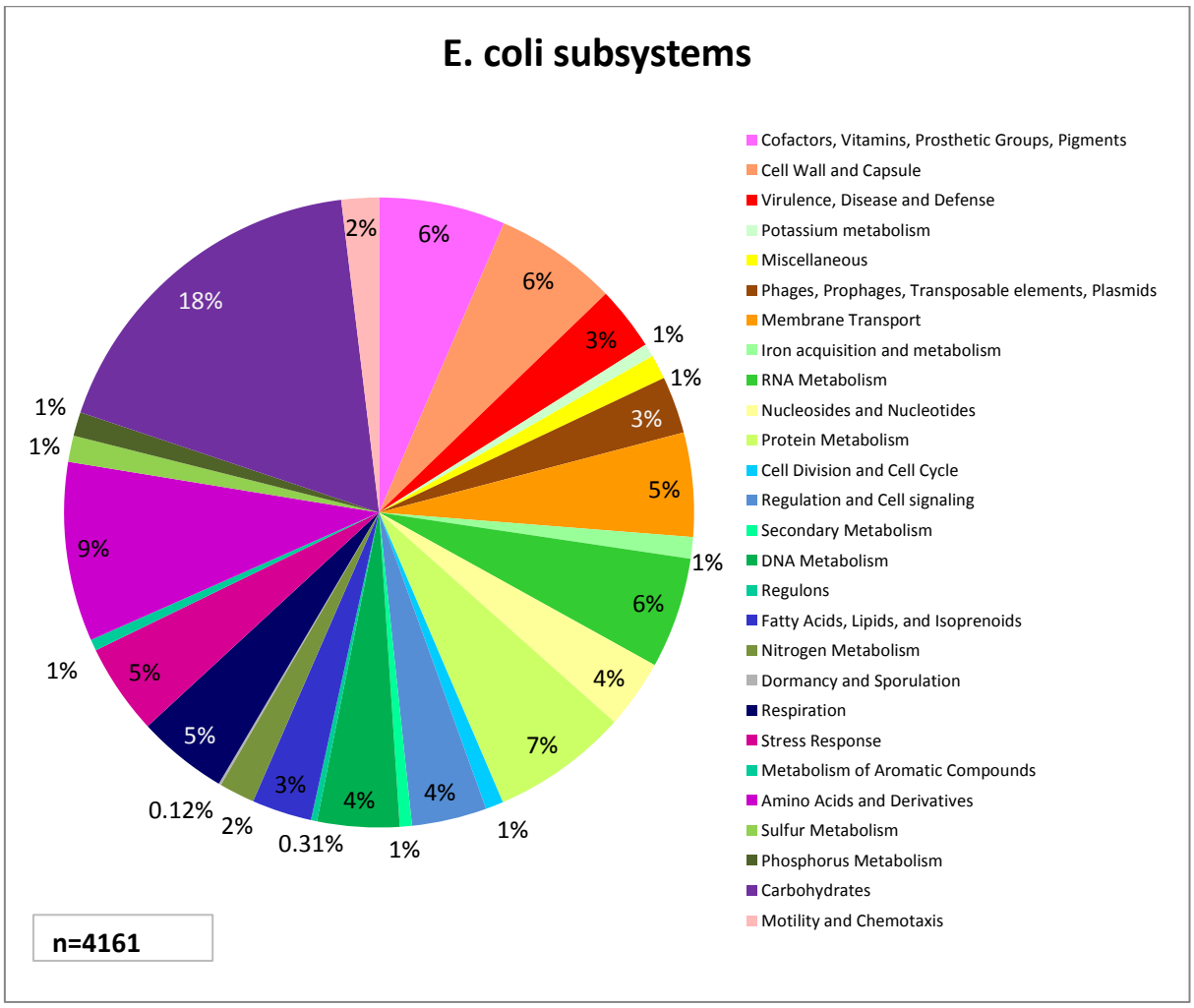


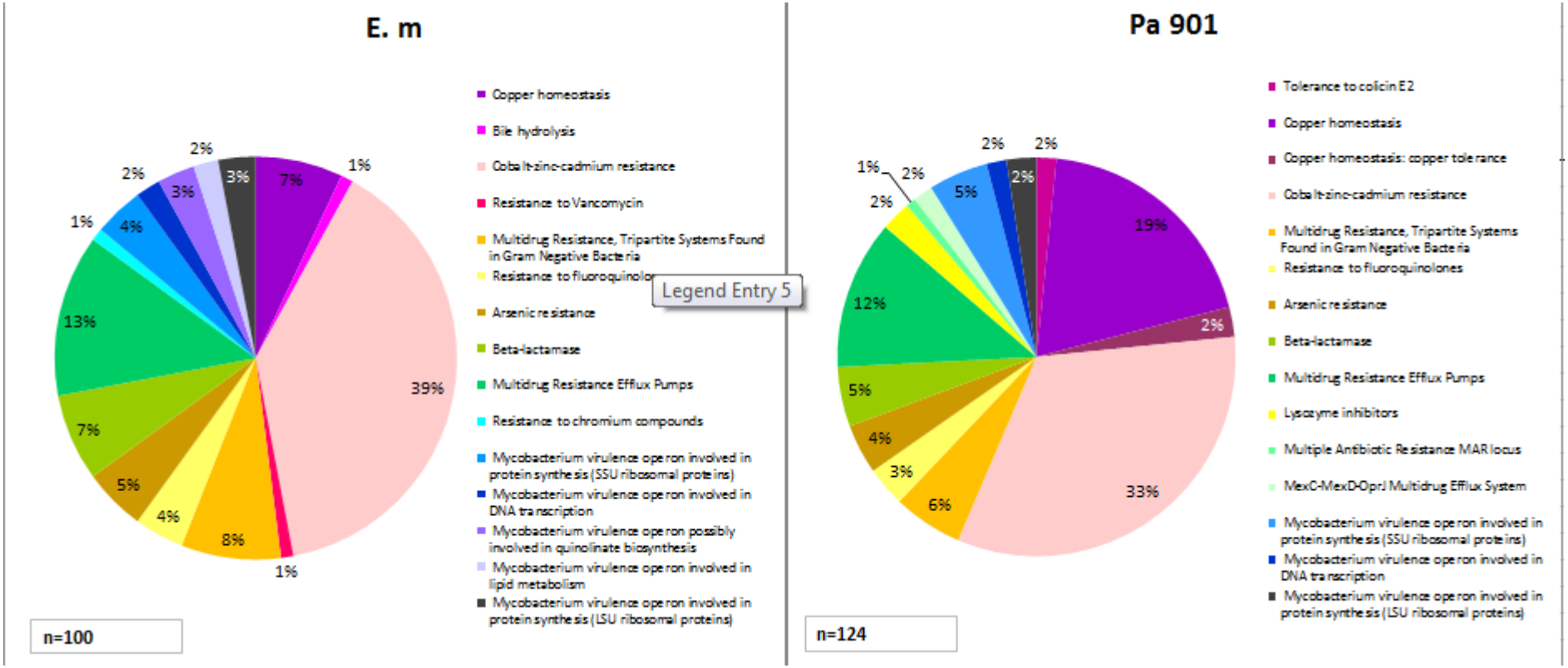
Chart 3.15. *E. coli* subsystems. Each subsystem is represented with different colour and its percentage ratio to the whole genome is showed. N – total number of features in subsystems.

3.4.3. Resistance patterns

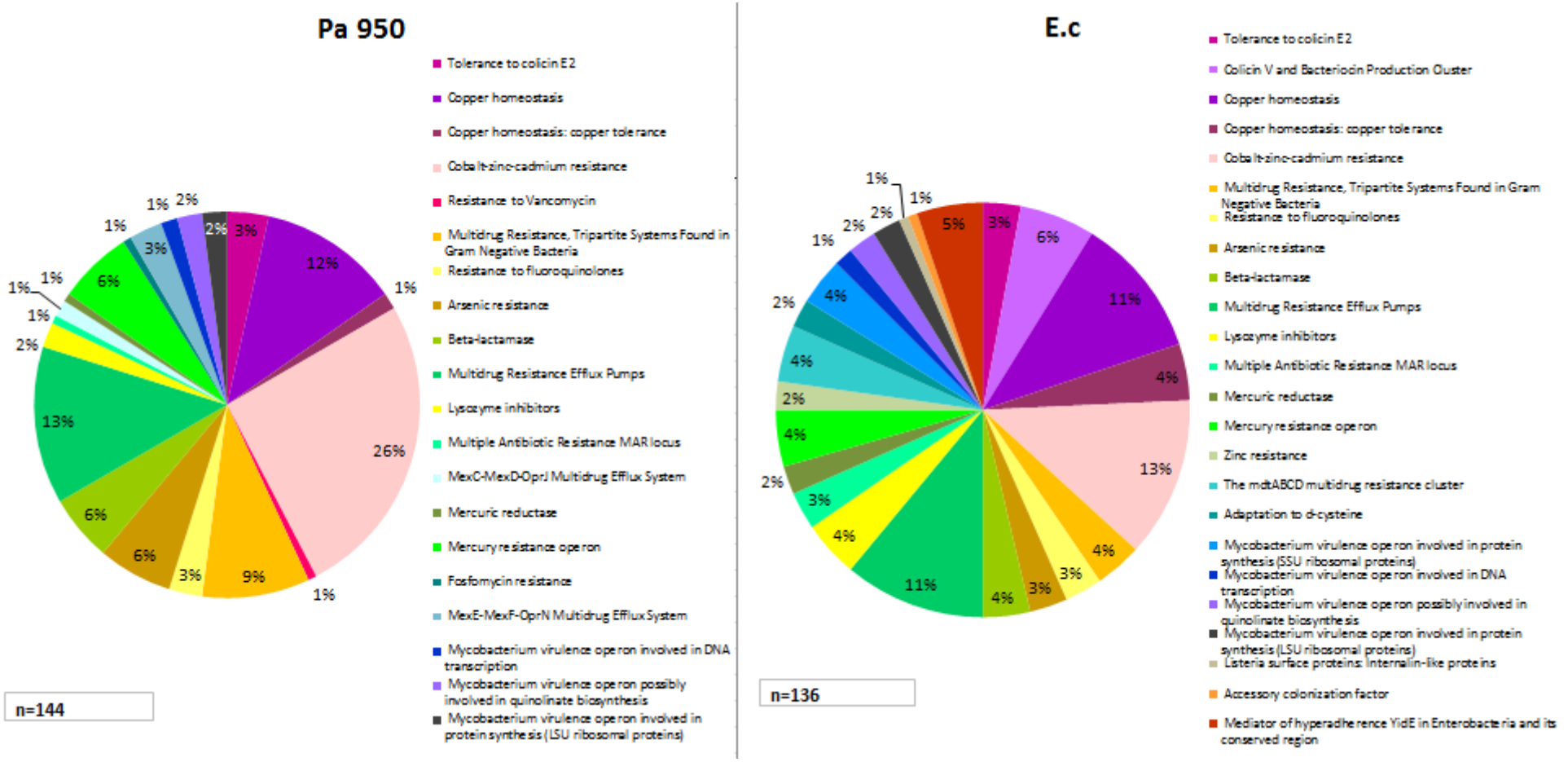
To prevent or treat the infection associated with pathogenic bacteria it is crucial to know the patterns of Virulence, disease and defence system. In order to survive bacteria have developed various systems, which became even more intricate in multidrug resistant pathogens. All four studied strains showed specific Virulence, disease and defence system patterns (**Chart 3.16**). The system was comprised of 4 categories: bacteriocins, ribosomally synthesized antibacterial peptides; resistance to antibiotics and toxic compounds; invasion and intracellular resistance; and adhesion (represented only in *E. coli*).

Bacteriocins and ribosomally synthesized antibacterial peptides (possessing tolerance to colicin E2) were represented in all bacteria, excluding *E. meningoseptica.* The largest and the most significant category – resistance to antibiotics and toxic compounds varied dramatically. The group, which represented cobalt-zinc-cadmium resistance, was the largest in all system varying from 13 (*E. coli*) – 39% (*E. meningoseptica*). Second biggest system's group (in *P. aeruginosa* 901, however, second biggest is copper homeostasis) – multidrug resistance efflux pumps, which are important as export different compounds (including antibiotics) from the cell were almost equally represented in all strains (11-13%). Other 5 groups, which were common for all 4 organisms comprised 3-19% of the system. All organisms were tolerant to copper, resistant to fluoroquinolones, arsenic and β-lactams, and possessed tripartite multidrug resistance systems. Opr efflux systems were specific for *P. aeruginosa* specimens and all strains, excluding *E. meningoseptica* had Multiple Antibiotic Resistance (MAR) locus, which play a role in efflux system regulation. Additionally, *E. meningoseptica* and *P. aeruginosa* 950 were resistant to vancomycin, but *P. aeruginosa* 950 and *E. coli* to mercury. Moreover, *E. meningoseptica* was resistant to chromium compounds, *P. aeruginosa* 950 to fosfomycin.

Furthermore, invasion and intracellular resistance category was comprised of virulence operons (involved in SSU and LSU protein synthesis, DNA transcription, quinolinate biosynthesis, lipid metabolism), which were fully represented in *E. meningoseptica*, and represented with some exceptions in *P. aeruginosa* 901, 950 and *E. coli*.


E. m
7%
1%
39%
1%
8%
4%
5%
7%
13%
1%
4%
2%
3%
3%
2%
Copper homeostasis
Bile hydrolysis
Cobalt-zinc-cadmium resistance
Resistance to Vancomycin
Multidrug Resistance, Tripartite Systems Found in Gram Negative Bacteria
Resistance to fluoroquinolones
Legend Entry 5
Arsenic resistance
Beta-lactamase
Multidrug Resistance Efflux Pumps
Resistance to chromium compounds
Mycobacterium virulence operon involved in protein synthesis (SSU ribosomal proteins)
Mycobacterium virulence operon involved in DNA transcription
Mycobacterium virulence operon possibly involved in quinolinate biosynthesis
Mycobacterium virulence operon involved in lipid metabolism
Mycobacterium virulence operon involved in protein synthesis (LSU ribosomal proteins)
n=100
Pa 901
19%
2%
33%
6%
3%
4%
5%
12%
2%
1%
2%
2%
5%
2%
2%
Tolerance to colicin E2
Copper homeostasis
Copper homeostasis: copper tolerance
Cobalt-zinc-cadmium resistance
Multidrug Resistance, Tripartite Systems Found in Gram Negative Bacteria
Resistance to fluoroquinolones
Arsenic resistance
Beta-lactamase
Multidrug Resistance Efflux Pumps
Lysozyme inhibitors
Multiple Antibiotic Resistance MAR locus
MexC-MexD-OprJ Multidrug Efflux System
Mycobacterium virulence operon involved in protein synthesis (SSU ribosomal proteins)
Mycobacterium virulence operon involved in DNA transcription
Mycobacterium virulence operon involved in protein synthesis (LSU ribosomal proteins)
n=124

a) **b)**



c) **d)**

Chart 3.16. Virulence, disease and defence subsystem's features in: a) *E. meningoseptica*, b) *P. aeruginosa* 901, c) *P. aeruginosa* 950 and d) *E. coli*. N – total number of features in subsystem.

3.4.3.1. *E. meningoseptica* resistance

The resistance of *E. meningoseptica* was studied in details and species with the most similar resistance patterns were identified. According to the **Chart 3.17** resistance patterns of *E. meningoseptica* most closely resemble *Flavobacterium johnsonia johnsoniae* UW101 patterns, sharing over 50% of chromosomal regions. Resistance of *E. meningoseptica* is also similar to *Chitinophaga pinensis* DSM 2588 and *Spirosoma linguale* DSM 74 resistance, sharing 36 and 26 % of chromosomal regions, respectively. Other species share less than 25% of resistance regions with *E. meningoseptica*. Interestingly, detailed analysis of *E. meningoseptica* resistance showed that bacterium shares several genes (4; example - **Chart 3.18.**) with *Anopheles gambiae* (African malaria mosquito).

No	Species sharing similar resistance chromosomal regions	Amount
1	Flavobacterium johnsonia johnsoniae UW101	56
2	Chitinophaga pinensis DSM 2588	36
3	Spirosoma linguale DSM 74	26
4	Pedobacter heparinus DSM 2366	21
5	Dyadobacter fermentans DSM 18053	19
6	Flavobacterium psychrophilum JIP02/86	18
7	Cytophaga hutchinsonii ATCC 33406	12
8	Croceibacter atlanticus HTCC2559	12
9	Gramella forsetii KT0803	8
10	Robiginitalea biformata HTCC2501	6

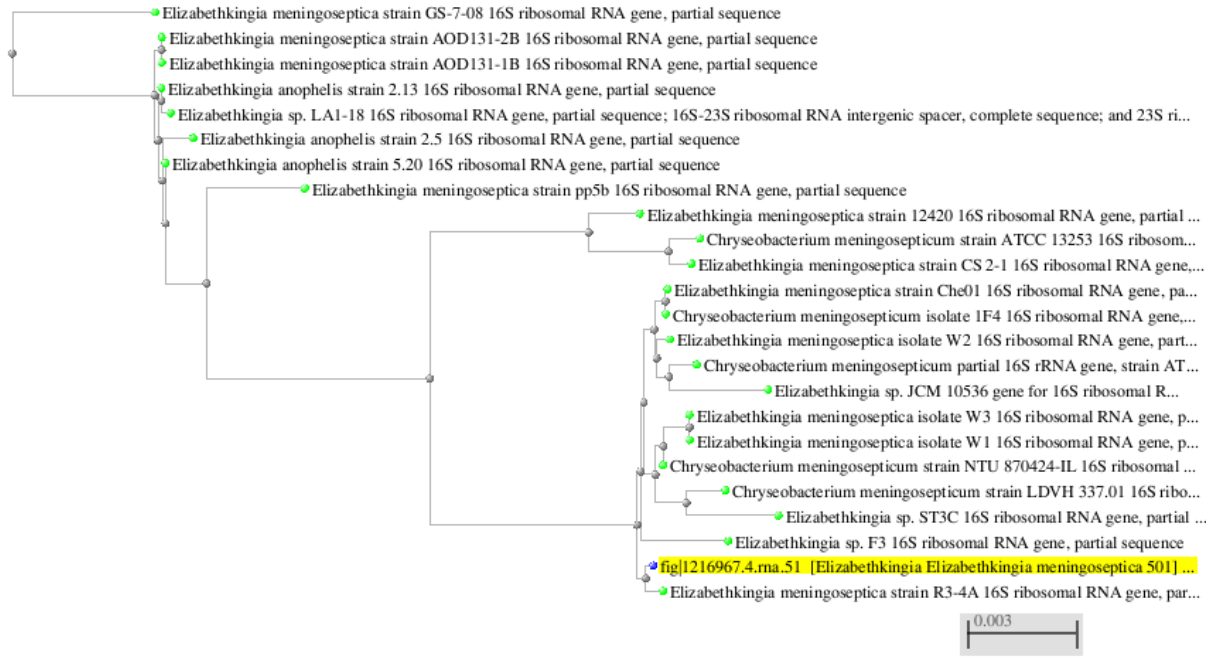
Chart 3.17. Species sharing similar patterns of resistance with *E. meningoseptica*. Numbers (right) represent the amount of chromosomal regions which are similar to *E. meningoseptica* (out of 100).

Resistance gene product	Organisms with similar chromosomal region	Contig No
Multicopper oxidase	Anopheles gambiae (African malaria mosquito)	Contig 10
	Anopheles gambiae str. PEST	
	Flavobacterium johnsonia johnsoniae UW101	
	Chitinophaga pinensis DSM 2588	

Chart 3.18. *E. meningoseptica* gene product (multicopper oxidase) comparison to other similar species (created by RAST).

3.4.4. Phylogenetics

Furthermore, relatedness of studied strains to other bacteria was analysed building phylogenetic trees in snpTree1.1 software (excluding *E. meningoseptica*). Due to the fact that no complete genomes of Elizabethkingia genus are available, *E. meningoseptica* trees were built using 16S Small Subunit Ribosomal RNA and DNA-directed RNA polymerase beta subunit (EC 2.7.7.6) in BLASTN software (**Figure 3.14**). According to 16S subunit tree the most closely related strain to *E. meningoseptica* 501 is *E. meningoseptica* R3-4A. DNA-directed RNA polymerase based tree showed complete bacterial genomes that are most closely related to 501 strain. From those *Riemerella anatipestifer* DSM 15868 and RA-GD strains appeared to be the closest relatives of *E. meningoseptica* 501 (also according to RAST).



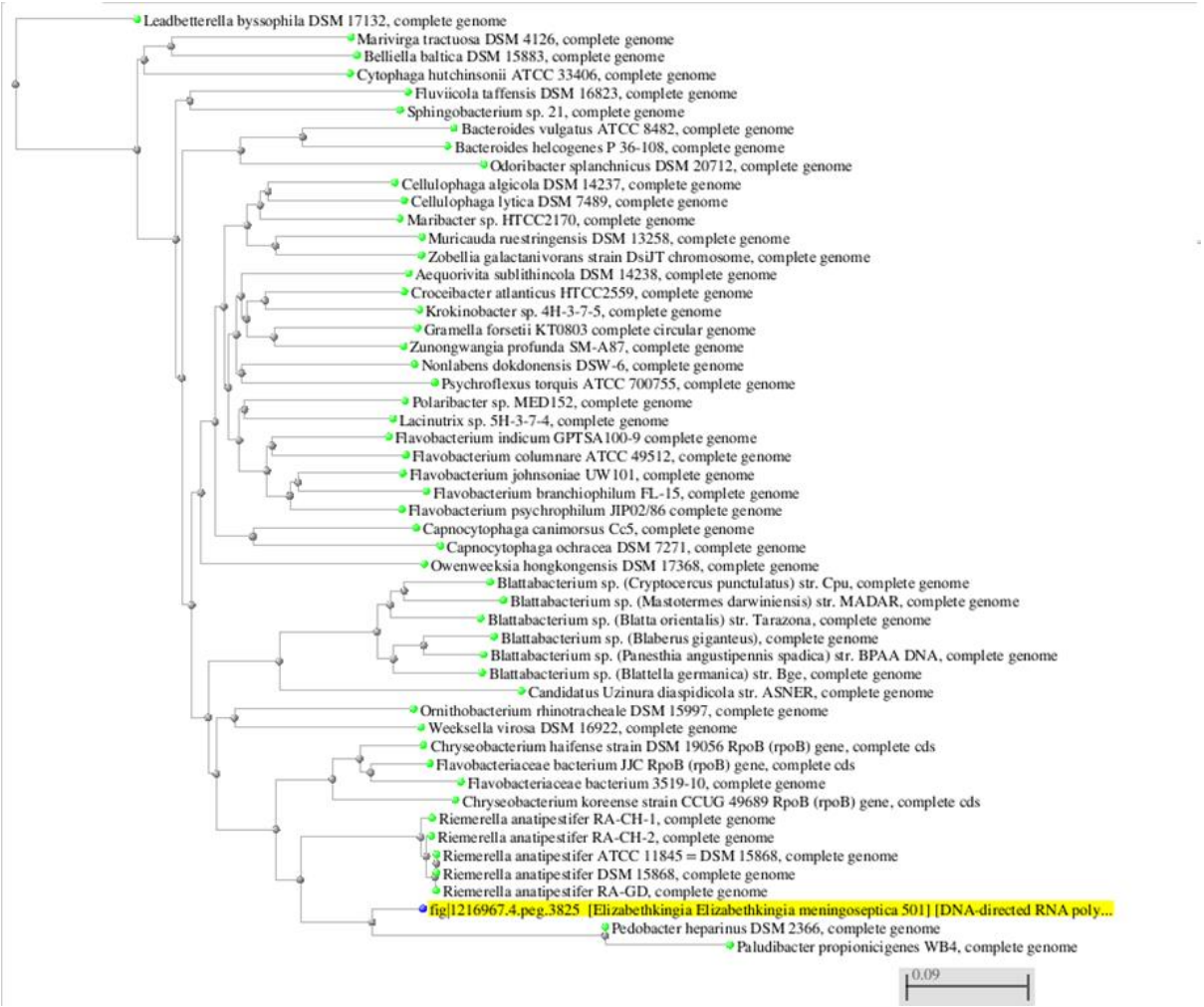


Figure 3.14. *E. meningoseptica* phylogenetic trees (strain 501 is highlighted). 16S Small Subunit Ribosomal RNA (top) and DNA-directed RNA polymerase beta subunit (bottom) based trees. Number at the bottom shows the evolutionary distance between strains.

P. aeruginosa phylogenetic tree (**Figure 3.15**) shows the relatedness of 901 and 950 specimens to all other *P. aeruginosa* strains (also showed in **Appendix IV**) with complete genomes (available from snpTree1.1). *P. aeruginosa* 901 and 950 are closely related as expected. Specimen 950 also is closely related to LESB58 strain. However, RAST software (using different approach – comparing the amount of same genes present in other species) showed that both specimens are most closely related to 19BR strain (not present in snpTree1.1 system) and NCGM2_S1 strain.

Nevertheless, *E. coli* phylogenetic tree (**Figure 3.16**) shows that from a variety of genomes available O104:H4_1218_280 strain is most closely related to strain 55989 (also **Appendix IV**).

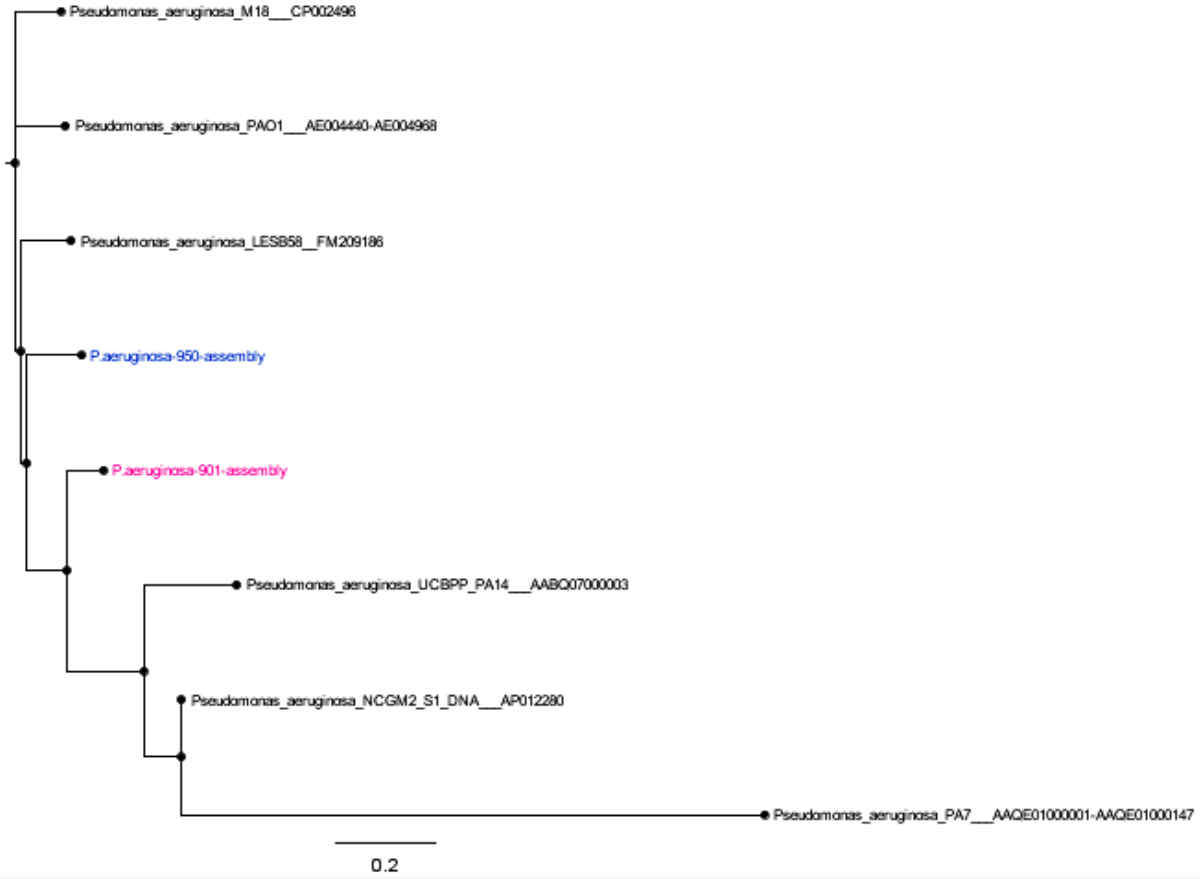


Figure 3.15. *P. aeruginosa* phylogenetic tree. Studied specimens (901 and 950) are coloured. Number at the bottom shows the evolutionary distance between strains.

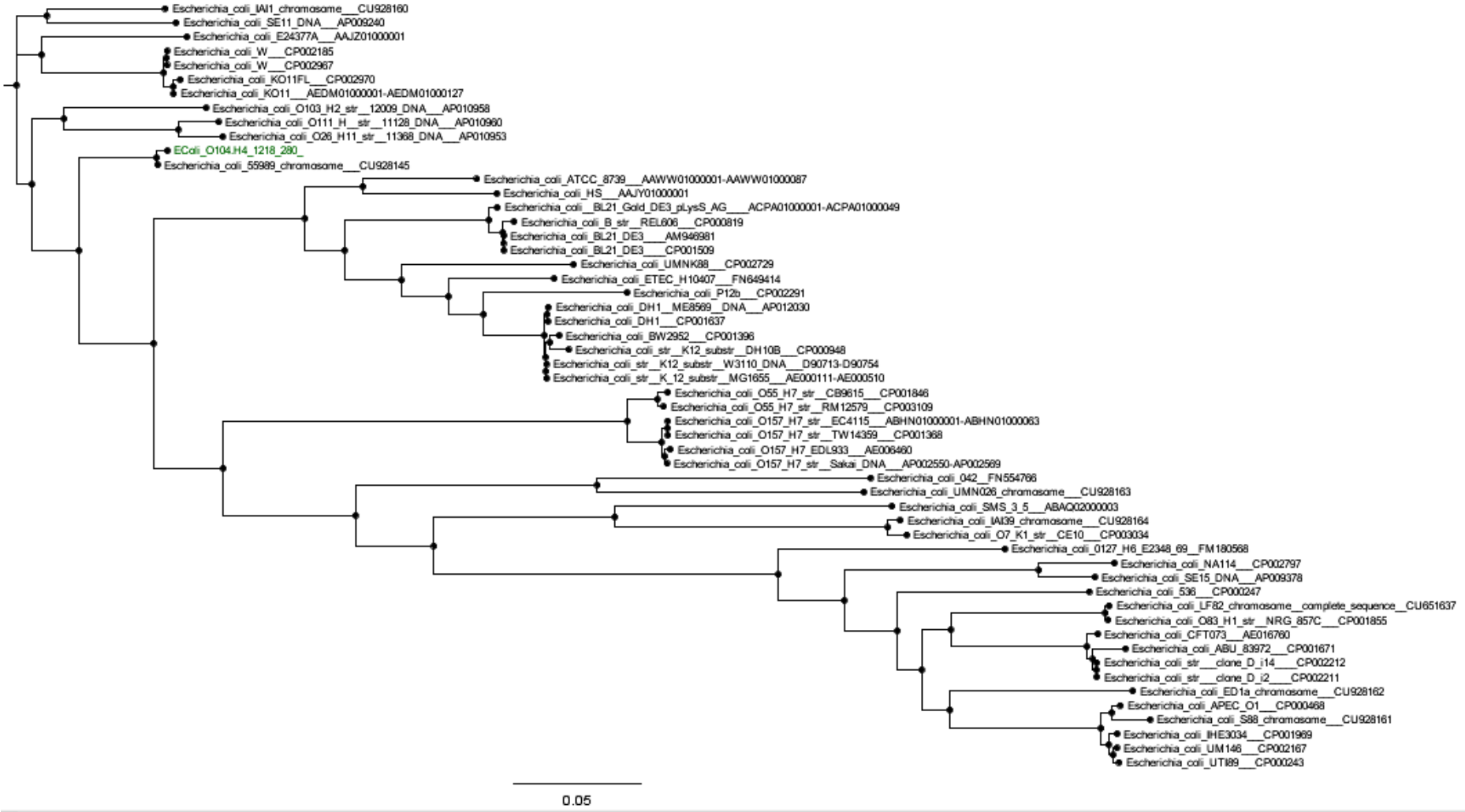


Figure 3.16. *E. coli* phylogenetic tree. Studied strain (O104:H4_1218_280) is coloured in green. Number at the bottom shows the evolutionary distance between strains.

4. Discussion

4.1. Paired-end whole-genome sequencing (Nextera XT)

The development of widely available and inexpensive high-throughput DNA sequencing has opened new horizons in life sciences, allowing the sequencing of large genomes in a matter of days. In microbiology HTS enabled discovering completely new approaches in many disciplines including gene expression analysis, mutation mapping, analysis of noncoding RNAs, metagenomics and whole-genome analysis (Lu et al., 2005; Malde, 2008; Mardis, 2008; Pop & Salzberg, 2008).

The accurate and successful preparation of DNA libraries for next-generation sequencing is a crucial step, which can be very challenging, and particularly for isolates with low DNA yield. Nextera XT Illumina sample preparation kit offers a solution for this problem as requires only 1 ng of input DNA. Nextera technology was designed in 2009 by Epicentre Biotechnologies (Caruccio et al., 2009), further purchased by Illumina Inc. in 2010. Nextera XT is one the most recent Nextera kits launched by Illumina in 2012. Due to the fact that technology is completely new only a few papers has been published employing Nextera paired-end strategy. Recently, Marine et al. (2011) evaluated Nextera protocol suitability for microorganisms using six phage genomes. The study indicates that the preciseness of Nextera sequencing was more than 99%, when comparing sequences obtained with Nextera and complete genome sequences available. Moreover, *De novo* genome assemblies, obtained in the same study, provided long contigs. These two factors make Nextera protocol a really good method for *de novo* assembling.

So far no study attempted to optimize Nextera XT protocol, therefore this study findings are novel. Optimization of sample preparation conditions showed that incubation time has no effect on the results, on contrary, DNA input concentration influences fragment length, but

only in *E. coli* samples. Nevertheless, this impact may increase in future experiments, if broader ranges of concentrations are used.

According to Marine et al. (2011), library fragments (mean) were 300 – 800 bp long, what is a good result. However, authors noticed that the mode value was higher in samples with higher DNA concentration, assuming that this is an effect of fragment 'nesting' – a phenomenon, when during limited-cycle PCR, along with the concentration increase, fragments create non-covalent bounds between the adaptors (it forms a false gel line – narrow and very concentrated – in Bioanalyzer results). Nevertheless, in current study no such an effect was observed. Marine et al. also indicated that GC content may affect the coverage of sequences, as during PCR step GC-rich and GC-poor sequences may be underrepresented (Aird et al., 2010; Kozarewa et al., 2009). Nevertheless, using the appropriate coverage Nextera is able to generate a high-quality data for many applications (Marine et al., 2011). In current study, however, no correlation between GC and fragment insert sizes was determined. The mean insert size values were similar for all species being 477, 526 and 521 base pairs for *Pseudomonas aeruginosa* 901 (high GC), *Escherichia coli* (medium GC) and *Elizabethkingia meningoseptica* (low GC) respectively.

4.2. Mate Pair sequencing

Mate Pair technology, as well as conventional paired-end method, apart from sequence data provides information about the distance between two reads. This information is valuable in structure rearrangement (e.g. insertions, deletions) determination and in assembling of repetitive regions. However, unlike paired-end data, Mate Pair data can provide user with information about the regions separated by large distances (2-20 kb; Nieuwerburgh et al., 2011). Mate pair technique generates 3 types of fragments: long range or mate pairs, short range or paired-ends and junction pairs, containing 1 chimeric read (only 1 end of the initial fragment instead of 2; Walenz et al., 2011). Obtaining the high amount of desirable high-

quality mate pair fragments is very challenging, as in many cases libraries are overrepresented by large quantities of paired-ends and junction pair fragments.

Nextera Mate Pair sample preparation kit was introduced by Illumina only half a year ago (January 2013), therefore no studies have evaluated the sample preparation efficiency yet. This new, improved (the first Illumina Mate Pair Library Prep V2 kit was introduced in 2009) Mate Pair kit has several important features: it is the only kit with gel-free method, which generates Mate Pair libraries up to 12 kb; gel-free method is very efficient, when DNA amounts are limited; preparation requires only 2 days (Illumina, 2013).

Until now only several studies employed Illumina Mate Pair technology to sequence, for instance, the genome of yeast (*Saccharomyces cerevisiae;* Nieuwerburgh et al., 2011) and human cancer cell lines (Arlt et al., 2011; Murphy et al., 2012). However, no study attempted to use it in microbiology. Nieuwerburgh et al. (2011) indicated the main disadvantages of Mate Pair (Library Prep V2, Illumina) protocol. First, authors note, that biotin enrichment step is not ideal, therefore longer sequencing reads result in probability of crossing the junction adapter (the longer read length the higher probability). Moreover, the junction adapter has no recognizable sequences, leading to difficulties in identifying reads that passed through the junction, and complicating mapping process. Consequently Library Prep V2 protocol suggests limiting the read length to only 36 bases (Nieuwerburgh et al., 2011). Nextera Mate Pair protocol offers a solution as contains identifiable junction adapter sequences, moreover, to minimize adapter bias problems Illumina provided a technical note (Illumina, 2012b) for computationally analysing and processing chimeric reads. Second problem with Library Prep V2 protocol is the large amount (18) of PCR cycles (Nieuwerburgh et al., 2011). PCR often produces duplicate sequences and amplification bias, what decreases density of libraries, therefore it is essential to minimize cycle amount when possible (Kozarewa et al., 2009). This problem was also eliminated with creation of Nextera Mate Pair protocol, drastically reducing PCR cycle number to only 10 (in gel-free protocol).

Present study indicates that even without optimization Nextera Mate Pair protocol allows generating a high-quality data, which undoubtedly facilitate genome finishing effort. Moreover, mate pair fragments are of a good length, and fragment distribution is very similar across all species, even despite that *P. aeruginosa* fragment yield is lower (however its genome is very complex as well). These facts indicate that protocol is robust and reliable. Nevertheless, it is reasonable to apply Mate Pair gel-plus protocol in future experiments, what most likely will give even better results.

Additionally, when preparing Mate Pair libraries it is also important to note GC content of an organism. In current study coverage (both in Nextera XT and Mate Pairs) increased when GC content decreased, however it presumably was due to the decreasing size of bacterial genomes (largest genome as well as GC in *P. aeruginosa,* and smallest genome and GC in *E. meningoseptica*). Therefore, most likely, there is no link between coverage and GC, though it requires further investigation.

Furthermore, library insert size (main) obtained in study by Nieuwerburgh et al. (2011; the authors used modified protocol with their-invented Cre-Lox recombination to improve Mate Pair performance) were around 3 kb, with the mode value ~ 2200 base pairs. In other studies by Arlt et al. (2011) and Murphy et al. (2012) similar results were obtained (main insert size ~ 3 kb). All studies, however, used Mate Pair Library Prep V2 protocol. Despite that no study tested the Nextera Mate Pair protocol yet, in the technical note (Illumina, 2012b) Illumina provided a model example of *E. coli de novo* genome assembly. The results indicated that median insert size was approximately 3 kb (mean not shown) and the mode ~ 2200 base pairs. Similarly, in current study mode insert sizes in all bacteria were close to 2000 bp, and the amount of high-quality mate pair fragments was promising, suggesting that the new protocol is working well.

4.3. *E. meningoseptica* genome finishing

The final step of whole-genome sequencing project, when sequences are already assembled into scaffolds, involves gap closure (captured gaps) and scaffold end joining (uncaptured gaps). The gaps are formed due to regions that are missing from contigs – unclonable sequences, GC-reach and GC-poor regions or due to large repeats (Tettelin et al., 1999). Before the development of Next-generation sequencing finishing process was laborious and time consuming, and to span numerous gaps additional steps were required. For instance, it was possible to construct a shotgun clone library employing plasmid vectors and then perform additional double-strand sequencing with universal forward and reverse primers. This approach generated sequence data from both ends of most clones (but not all), and was followed by primer walking either on shotgun clone template (Fleischmann et al., 1995) or genomic DNA (Heiner et al., 1998). The method became more laborious, when the gap was larger than few hundred bp, moreover, the obtained connections required confirmation by PCR. Furthermore, the presence of repetitive regions made the process even more complicated as the average sequence read (of Sanger sequencer) was smaller than most of repeats, and the primer located outside could not get across the repeat. Therefore, one more additional PCR step was required to generate product across the gap, and perform primer walking directly on PCR product. However, this process required to test each contig end against all of the other ends, resulting in hundreds or thousands different combinations and PCR reactions (multiplex PCR (Burgart et al., 1992) and combinatorial PCR facilitated this step; Tettelin et al., 1999).

Nowadays, HTS technology significantly simplifies scaffolding (especially after Mate Pair sequencing) and finishing effort making the number of possible connection variants minimal. To close the gaps in-between contigs, which are joined into scaffolds (captured gaps), is very easy as the connection between 2 contigs as well as direction of the contigs (flipped or non-

flipped) are already known. Scaffold end joining (uncaptured gaps) is more complex as few combinations are still possible (which are determined by such programs as Gap5 and Tablet), however, due to the development of HTS technology testing each contig end against all of the other ends is unnecessary. Finishing step after high-throughput sequencing requires only 2 steps – PCR with the unique primers, which are close to the end of the contig and PCR product sequence determination.

If the amount of uncaptured gaps is still high, various PCR approaches can be used instead, including multiplex PCR (Chamberlain et al., 1988), inverse PCR (Triglia et al., 1988), restriction site PCR (Sarkar et al., 1993), capture PCR (Lagerstrom et al., 1991) and anchored PCR method (Mizobuchi & Frohman, 1993; Siebert et al., 1995; Rogers et al., 2005).

When the number of uncaptured gaps is low they can be closed relatively fast by the direct PCR with every possible pair-wise combination of primers. After PCR product is obtained Sanger sequencing is performed to determine the gap sequence. This technology was successfully used in current study, and in a very short time period (~ 2 month) from initial 50 contigs of *E. meningoseptica* genome 4 scaffolds were finally obtained. However, some contig connections cannot be determined using the original Taq polymerase as contain large repeats (6 or more kb) in-between contigs, therefore these few products should be obtained using Long Range PCR. Therefore, in order to complete *E. meningoseptica* genome some additional experiments (similar and not time-consuming) are still required. Furthermore, it is important to note that the progress in sequencing technology already greatly facilitated the genome finishing process, and further development of paired-end sequencing will most likely diminish the role of PCR in finishing process even more (Nagarajan, 2010).

4.4. Genome annotation

So far no study addressed genome features of any Elizabethkingia genus members. In this study general features of *E. meningoseptica* genome were described and the approximate genome size was estimated for the first time. Being ~ 4.2 Mb long it is smaller than genomes of *P. aeruginosa* or *E. coli*, however still large genome for bacteria. The genome of *E. meningoseptica* close neighbour *Riemerella anatipestifer* DSM 15868, for example, is twice smaller (2.2 Mb; Yuan et al., 2011). *E. meningoseptica* resistance patterns also seem less complex and diversified than patterns of 2 other bacteria studied, especially due to the fact that this bacterium has no plasmid. Still quite large part of its genome (5%) is responsible for resistance. Additionally, interesting fact that bacterium shares several genes with *Anopheles gambiae*, may be explained by the inclusion of *Elizabethkingia anopheles* DNA fragments (which are highly similar to *E. meningoseptica*) into the genome of *A. gambiae*. Nevertheless, further analysis of *E. meningoseptica* genome is required to obtain a complete resistance picture.

Obtaining a complete genome of *P. aeruginosa* is a great challenge due to the complexity and size of its genome. There are more than 10 *P. aeruginosa* genomes finished so far and 9 of them are available as a single chromosome from NCBI website database. As indicated by RAST system, specimen 950 genome and resistance patterns are more complex than 901, what is most likely due to the fact that 901 was sampled in water and 950 persisted in human. Strain 19BR is the closest relative of both specimen 901 and 950. This strain is multidrug resistant, and apart from usual fluoroquinolones, β-lactams resistance caries specific resistance to polymyxin B (widely used drug against Gram-negative bacteria; Boyle et al., 2012). Second closely related strain NCGM_S1 induced urinary tract infection outbreak in Japan and were extremely resistant to aminoglycosides and β-lactams (Miyoshi-Akiyama et al., 2011). Another related strain – LESB58, (originated from Liverpool) that cause cystic

fibrosis in children, is specifically known for its ceftazidime resistance (Winstanley et al., 2009). Furthermore, future experiments are required to complete the genome of *P. aeruginosa* ST395 and compare in details its genomic features to other genomes.

Escherichia coli O104:H4 strain was extensively studied during 2011 outbreak of hemolytic uremic syndrome in Germany, thereby it is well-characterized. Complete genome sequence of German isolate is already published and the genome is annotated (Ahmed et al., 2012). Genome size, number of plasmids (3) and GC content of German isolate is very similar to isolate 1218_280 sequenced in current study. O104:H4 strain belongs to enteroaggregative type and is characterized by Shiga toxin 2 production and tellurite resistance. Both Shiga toxin 2 and tellurite resistance cluster are also present in O104:H4 1218_280 specimen. Additionally, strain 55989 was determined to be the closest O104:H4 relative (Rohde et al., 2011), what is in agreement with current study.

5. Conclusions

The whole-genome sequencing (paired-end and mate paired) of three different bacteria - *E. meningoseptica, P. aeruginosa* and *E. coli,* was performed in this study. Nextera XT protocol optimization step one showed no significant concentration or time-dependant trend, however yielded fragments over 450 bp (mean) for all three bacteria, which suggest of a high efficiency of the protocol. During the second Nextera XT optimization step significant concentration-dependant trend was observed in *E. coli* samples (other bacteria showed no trend), therefore it was concluded that concentration of 0.4 ng will be most suitable in further sequencing steps. Paired-end sequencing also allowed joining sequences into contigs and created high-quality draft genomes of bacteria. Nextera Mate Pair sequencing considerably improved the results – the insert sizes as well as coverage were much greater than in paired-end sequencing. Mate Pair information also allowed joining contigs into scaffolds and estimated the gaps and direction of contigs within scaffolds, what considerably simplified finishing effort. For the last step of genome finishing (only for *E. meningoseptica*), PCR was successfully optimized, and given above conditions are considered to be more efficient for finishing than manufacturer's protocol. However, Long Range PCR protocol still requires optimization. PCR step enabled joining scaffolds of *E. meningoseptica* into 4 (gaped) fragments only after several months of continuous experiments, what demonstrates an enormous progress of high-throughput sequencing during the last decade. Furthermore, genome features, such as genome size, subsystem distribution and patterns, resistance patterns, phylogenetics were discussed in the study. The genome of *E. meningoseptica* was described for the first time and it is the smallest and most likely the simplest genome amongst bacteria featured in current study. However, the resistance patterns of *E. meningoseptica* are still complex and resistance subsystem is a considerable part of its genome. Resistance

patterns of *P. aeruginosa* 950 and *E. coli* are the most diversified amongst studied bacteria, containing a wide range of resistance systems. Nevertheless, phylogenetic analysis showed that *Riemerella anatipestifer* appears to be closest *E. meningoseptica* neighbour. Both *P. aeruginosa* specimens appear to be close neighbours of 19BR strain, and *E. coli* is closely related to 55989 strain. In summary, modern technology already greatly facilitated finishing of small bacterial genomes and with further development HTS technology may be successfully used in many different applications.

Appendices

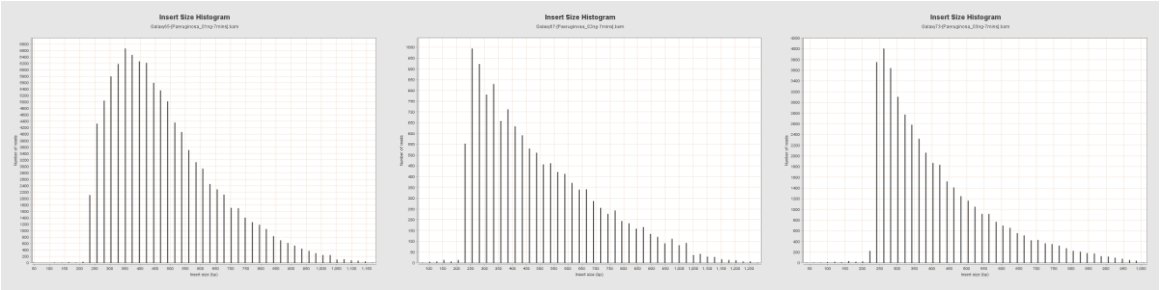
Appendix I

Contig No	5'FORWARD primer	3'REVERSE primer	5'REVERSE primer	3'FORWARD primer
EM_S22_c1	TGCAATAGCTGTACGCATCA	ACTCCCCGCACAATACAGTC	GGCAGAAACCTTCTCTAGCAA	CCAACATGGCTTACGGAAAA
EM_S22_c2	GGTTAAAACGCAGGTTGAGG	CGTTTTGCCATCTTAGTGAACTT	CCCGGCCATATTGAAAAGAT	CGGCTACCCGGTAATTGTAA
EM_S22_c3	CCCTCCAAGATCTGCAACAT	AACGCCGTGCTTAGAGGTTA	CAGGCTTTGGAATAGCAACC	AAAACCCTTTTGCCGGATT
EM_S22_c4	TTTATCACGTGGAGCAGCAG	AGCTCTAATCCGTGAGAATCG	TTCATATACTCCGAATTTCTGGAAC	TTTTGAGGATAACCATAAACAAGTAA
EM_S22_c5	GCAGACTTGCCTCCTGTTCT	CCGTGTCTTTATTGCTCTTGC	AGGCACTTCCGTTATCGTTG	GGTAATTTTCCAGCTTTACTTTTATCA
EM_S22_c6	ACCAAGTGAAAATGGAAGCTTTA	TGGGGAAATTCCAAAACAAA	AAACAAATTAAAGCCCCAAAAA	TGATATCGACAGAATGCAGGA
EM_S22_c7	GAATTCCGAGCACCGATAAA	AGACCCGGATACAGGACCTT	GTGCTCGGAATTCAGGATGT	ACATGGATCATTTCCGGTTT
EM_S22_c8	AGGTTGCTCACCTGCCATAA	AGTAGGTCAGCAGCCTTTCG	CCGTAGATAATTCTGGCGAAA	GGAAATCCGAGGCTCTCATA
EM_S22_c9	GGGAAACCGTAATGCTGTGT	CAGCGCTTATTGTTGCAGAA	GCAAAAGCCTGAAGTCGTAGA	TTTCCTCTGTACTCAGGCTTTACA
EM_S22_c10	GATTGGTGACGCCAAAACTT	AAGAAAACCGCGTATCATGG	AAGTTTTGGCGTCACCAATC	TGGTGTTTATATGCTGTGCATT
EM_S22_c11	AGAACTCCCGCTACAACAGC	GACCGATAAACCTTACGAGCTT	CTCCGTTTCTGGTATTGTTGG	TGGTTTGAAGCCGGAGTTAC
EM_S22_c12	TCCACGCAGGATATCCATTT	GATGGGGTACGAGGCTATCA	TACGCTATTTTTCCGGCATT	TGATAGCCTCGTACCCCATC
EM_S22_c13	TGGTGGTTTGGTGTAACGAG	GGAAATACTAAATGAACCGGAGTTT	TGCTATAGGGACCTCACTTGC	GAAAGATTGACCCGGGTTGT
EM_S22_c14	GATCGTTTCCAGCCAAACTC	CGGAAAAGTTTCGGATACCA	TTTACCTCATAGCTGGATTTCAA	TTGATTTTGGATAAATCTGTGTTGA
EM_S22_c15	ATTGTCGCTCCGGAATAAGA	TTGGAAGCAATAGCCTTCTTT	GCTGACACCAAGTTGCAAAA	TCCCGTTTTCAAAGTATCTGC
EM_S22_c16	GTGCTTCTGGAGGAGCAGAG	TTGCTTTCTTACGATAGAGTCAACA	TCGAGTTGTATCCCTGATGTTTT	AATGATGAAAACCGTTTCTTACAA
EM_S22_c17	GATTGCATACCCAGGAGGTG	CCTCCGGCTTCTAAAACACTC	AGTGGGCATCAGCATCTTTT	CCTTTATTCCTGCCGAAGGT
EM_S22_c18	TCAACAAATTGCAAGAAAGCA	TGTTTGCCTGATTTAAAATGACT	TCAGCTTTTTCAAAATTAGGAGGT	TGCATAACTCATTACCGATGCT
EM_S22_c19	TTTGTCATTGCTGCAAGTCC	TACCGACTTCGGAGGGGTAT	CATCTTCACGGAACTGAGCA	ATTCTTGCCCCGTAATCGTA
EM_S22_c20	CCGACGGAAATGGTCTTACT	TTCTCCGCTTTGAAAACGTC	GGCCATTAATTCTGAGAGATGG	GCATGGGCTCTTATGGCTAA
EM_S22_c21	TGTCAGGCAAGGATTTAAAATTG	TTTACGGAGCAAGAGCCATT	GGGAACAGGGTTTGGTCTTT	TCTTTTTGCAACTTTGGCATC
EM_S22_c22	TCGGAATATTTTTGGGTGGA	AAGTTGAACGGCGTTACTGC	ACTAAGCAATGCGCTCCAGT	GCTGAGATTGATGCAGCTTG
EM_S22_c23	TACTGCACAGCCAGGACAAC	CCTTAGGTAAAGCCCGGAAA	CCATACCAATCGTGTGCATT	TTCTTTGTCTGCGCTTTTCA
EM_S22_c24	AGTGCTGCAGGAAGCATTTT	CTTTGCTTCCCCTCTCTGC	TGGGACAAAATACCAGCTCTC	GGGAACTATATCGGCACAGG

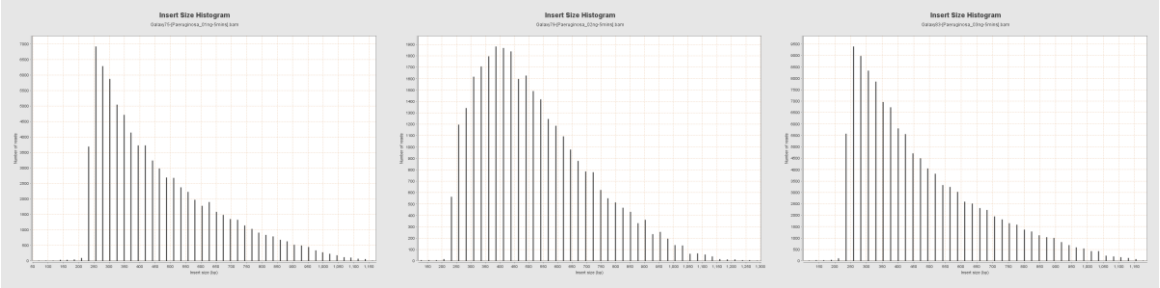
EM_S22_c25	TGGAAAGACTCCCCCTCTTT	AAATGCACATTCGCATAATGA	GCGTAAAGCCGGATTTTTCT	TCGAGGTAAAGGCTTTCCAA
EM_S22_c26	ATCCCCCACCAGTCTACCTC	CTCCAATCCAAGACCTTTGC	GCTTGGGATAATTGCAGCAC	GGGATGCTGAAGCTTGGTTA
EM_S22_c27	TTGCTAATCTGATGGTTGAATGTT	GCCAATAAAGAACGGGAGAA	CCGTCTTTACTTTTCAACATTGG	CACAATATGACCTTTCAGTACAAAGC
EM_S22_c28	TTCCTGGGTCATGATTTGGT	CGAATGTACCAGCTGTAACAAAG	GCCGGAACTAAGGTCAACAA	ATGGTTATCCGCCAGAATTG
EM_S22_c29	TCACTCCAACCTCCCAAAAC	TGCGTAATACACACTGGCTGT	ACGTCTCCGTTTCGGAATTT	GCGGGTTTTATTTTTGCTTTT
EM_S22_c30	AGCTTTTACGGGGGACAGTT	TGGACAGAAAGCAGCACGTA	CAACAGGGCCAACACCTTTA	CTGGTATCAGCTCCCAGGTT
EM_S22_c31	GAAGCTCCGGAAGAGTCTGA	AACTGAATACACGCGAGAAAAA	TCCAGAGGAATACTTACTTTCTGACA	GCCATTTTTATGATTTATAGGATGA
EM_S22_c33	CAGAAAAATTCCCTTTGGTCTT	CCTCCAAAAATTGACATGTTTC	TTGGAGCTGCAACAACTACG	TTCACCAGAATTATCTACTGCATTG
EM_S22_c45	GCCGACGGTGTATGGTAAAT	TGCAGCCACAATGCTTTATC	GCATCCAATCGGAAAACATA	AAATTGATAAAGCCCTCATTTG

Chart A.1. Primer sequences designed for *Elizabethkingia meningoseptica* 501 genome finishing. Unique primers of four different directions (3' and 5', forward and reverse) were designed for each contig (at least 500 bp from the end of the contig) employing Primer3 (v. 0.4.0) software.

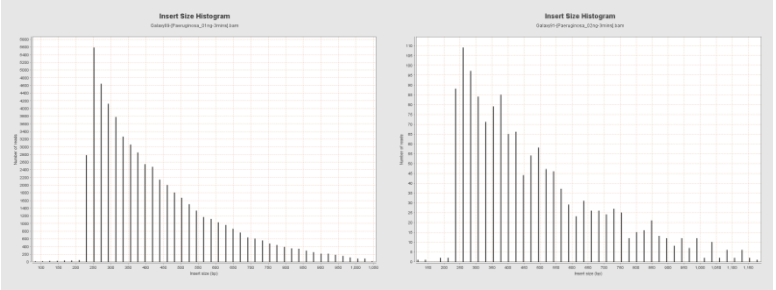
Appendix II.



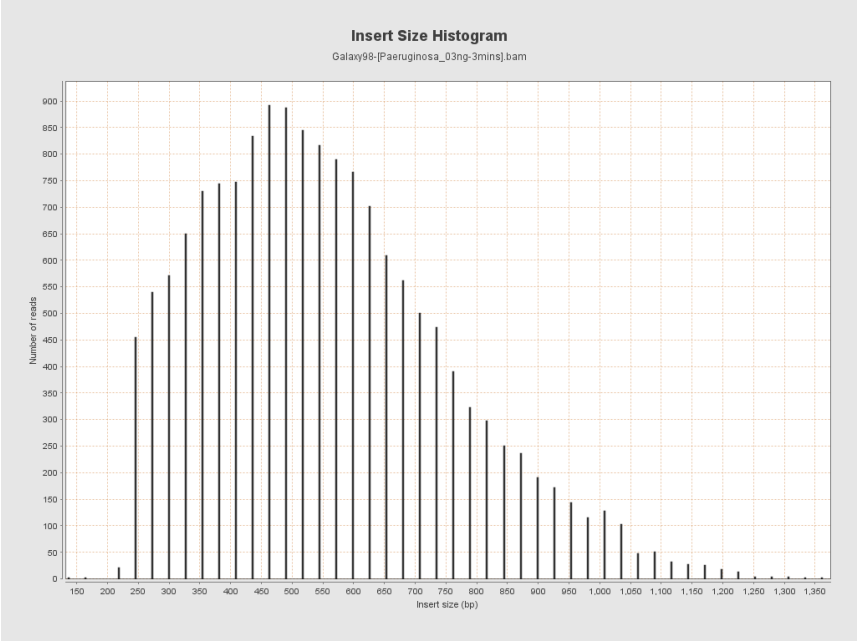
1) 2) 3)



4) 5) 6)

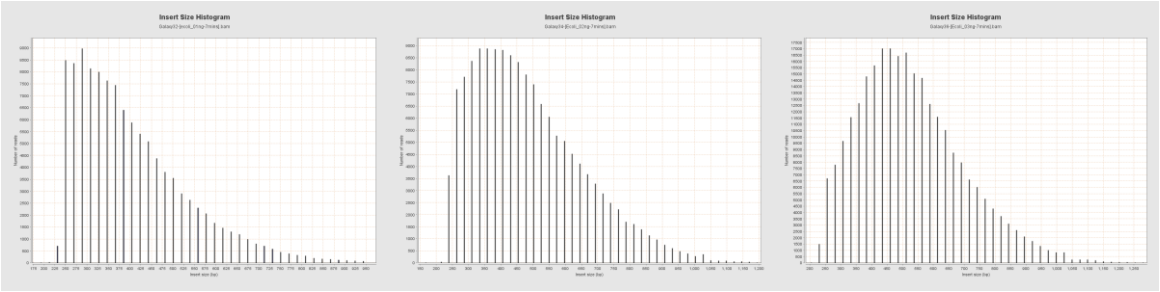


7) 8)

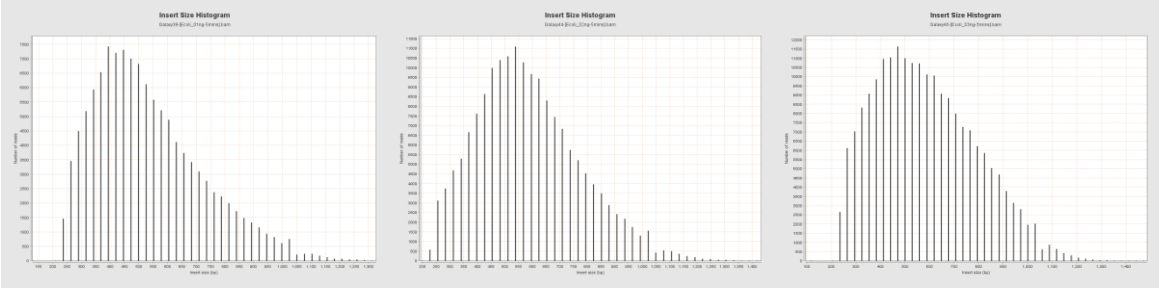


9)

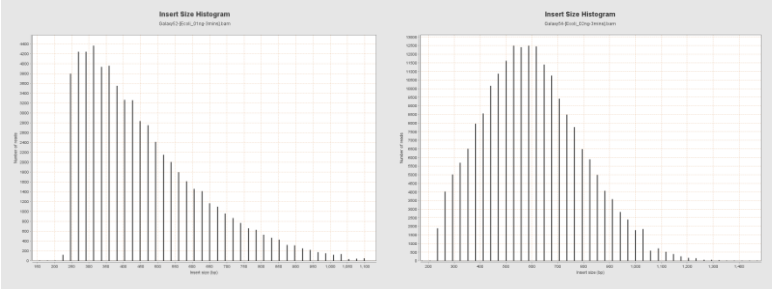
Chart A.2.1. Insert size histograms of *P. aeruginosa*. Bars represent number of reads at different insert sizes, bp. 1-9 show different sample preparation conditions (listed in Chart 3.1, column 2). Created using Qualimap v0.7.1.



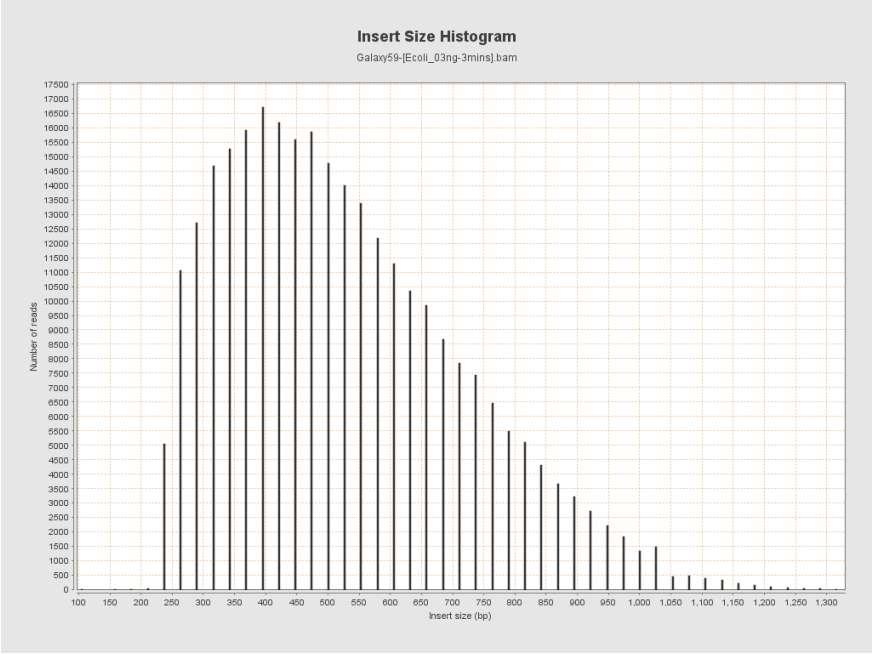
1) 2) 3)



4) 5) 6)



7) 8)



9)

Chart A.2.2. Insert size histograms of *E. coli*. Bars represent number of reads at different insert sizes, bp. 1-9 show different sample preparation conditions (listed in Chart 3.1, column 1). Created using Qualimap v0.7.1.

Appendix III.

Reactions	Primer		Primer		Reactions	Primer		Primer	
Reaction No	3'F	Repeat	5'R	Product size	Reaction No	3'F	Repeat	5'R	Product size
1	C5		C2	1,5 kb	19	C10	repC251	C16	2,75 kb
2	C2		C12	1,25 kb	20	C16		C33	1 kb
3	C12	rep C265	C21	2,25 kb	21	C33		C18	1,5 kb
4	C21		C28	1,75 kb	22	C9		C6	1,75 kb
5	C28		C4	1,5 kb	~~23~~	~~C7~~	-	~~C3~~	
6	C4	repC250	C25	-	24	C7		C30	900 bp
7	C25		C20	1,25 kb	~~25~~	~~C11~~	-	~~C17~~	-
8	C20		C7	2 kb	~~26~~	~~C11~~	-	~~C31~~	-
9	C19		C27	1,75 kb	27	C30		C3	700 bp
10	C27		C11	1,25 kb	28	C14		C15	2,5 kb
11	C8		C22	1,25 kb	~~29~~	~~C15~~	-	~~C13~~	-
12	C17		C24	-	~~30~~	~~C18~~	-	~~C17~~	-
13	C24		C29	1,25 kb	~~31~~	~~C15~~	-	~~C17~~	-
14	C29	repC256	C23	1,5 kb	~~32~~	~~C6~~	-	~~C17~~	-
15	C23		C26	1,25 kb	~~33~~	~~C31~~	-	~~C17~~	-
16	C26		C14	1,25 kb	34	C18		C31	-
17	C45		C1	550 b	~~35~~	~~C15~~	-	~~C31~~	-
18	C1		C10	-	36	C6		C31	-

Chart A.3.1. PCR reactions between different contigs, initial set. Created, based on Contig map No 1 and 2. Green colour indicates scaffolded connections, blue – PCR-determined connections, red – possible connections (probably containing large repeat, therefore product failed to develop), grey – excluded connections.

Reactions	Primer		Primer		Reactions	Primer		Primer	
Reaction No	3'	Repeat	5'	Product size	Reaction No	3'	Repeat	5'	Product size
37	C4_3'F		C25_5'F	-	38	C4_3'R		C25_5'R	-
~~39~~	~~C17_3'F~~	-	~~C24_5'F~~	-	40	C17_3'R		C24_5R	1,25 kb
41	1_3'R		C10_5'R	-	42	1_3'F		C10_5F	-
~~43~~	~~11_3'F~~	-	~~C17_5'F~~	-	~~44~~	~~11_3'R~~	-	~~C17_5R~~	-
~~45~~	~~11_3'R~~	-	~~C31_5'R~~	-	~~46~~	~~11_3'F~~	-	~~C31_5F~~	-
47	15_3'F		C13_5'F	550 bp	~~48~~	~~15_3'R~~	-	~~C13_5'R~~	-
~~49~~	~~15_3'F~~	-	~~C17_5'F~~	-	~~50~~	~~15_3'R~~	-	~~C17_5'R~~	-
~~51~~	~~18_3'F~~	-	~~C17_5'F~~	-	~~52~~	~~18_3'R~~	-	~~C17_5'R~~	-
~~53~~	~~6_3'F~~	-	~~C17_5'F~~	-	~~54~~	~~6_3'R~~	-	~~C17_5'R~~	-
~~55~~	~~31_3'F~~	-	~~C17_5'F~~	-	56	31_3'R		C17_5'R	500 bp
57	18_3'F		C31_5'F	-	58	18_3'R		C31_5'R	-
~~59~~	~~15_3'F~~	-	~~C31_5'F~~	-	~~60~~	~~15_3'R~~	-	~~C31_5'R~~	-
61	6_3'F		C31_5'F	-	62	6_3'R		C31_5'R	-

Chart A.3.2. PCR reactions between different contigs, set 2. Created, based on Contig map No 2. Green colour indicates scaffolded connections, blue – PCR-determined connections, red – possible connections and grey – excluded connections.

Reactions	Primer		Primer		Reactions	Primer		Primer		Reactions	Primer		Primer	
Reaction No	3'F	Repeat	5'R	Product size	Reaction No	3'F	Repeat	5'F	Product	Reaction No	3'R	Repeat	5'R	Product size
~~63~~	~~C22~~	-	~~C19~~	-	~~64~~	~~C22~~	-	~~C19~~	-	~~65~~	~~C22~~	-	~~C19~~	
~~66~~	~~C3~~	-	~~C9~~	-	~~67~~	~~C3~~	-	~~C9~~	-	~~68~~	~~C3~~	-	~~C9~~	-
69	C3		C45	-	70	C3		C45	-	71	C3		C45	-
~~72~~	~~C6~~	-	~~C19~~		~~73~~	~~C6~~	-	~~C19~~	-	~~74~~	~~C6~~	-	~~C19~~	-
75	C13		C19	3 kb	~~76~~	~~C13~~	-	~~C19~~	-	~~77~~	~~C13~~	-	~~C19~~	-
~~78~~	~~C11~~	~~repC250~~	~~C9~~	-	~~79~~	~~C11~~	~~repC250~~	~~C9~~	-	80	C11	repC250	C9	400 bp
81	C5_5'	repC254	C8_5'		82	C5_5'	repC254	C8_5'		83	C5_5'	repC254	C8_5'	
84	C22_3'		C3_3'		85	C22_3'		C3_3'		86	C22_3'		C3_3'	
~~87~~	~~C22_3'~~	-	~~C13_3'~~	-	~~88~~	~~C22_3'~~	-	~~C13_3'~~	-	~~89~~	~~C22_3'~~	-	~~C13_3'~~	
~~90~~	~~C11_3'~~	~~repC250~~	~~C6_3'~~	-	~~91~~	~~C11_3'~~	~~repC250~~	~~C6_3'~~	-	~~92~~	~~C11_3'~~	~~repC250~~	~~C6_3'~~	

Chart A.3.3. PCR reactions between different contigs, set 3. Created, based on Contig map No 2. Blue colour indicates PCR-determined connections, red – possible connections and grey – excluded connections.

Appendix IV.

Strain	Size, Mb	GC,%
Pseudomonas aeruginosa ST395 901	5.94	66
Pseudomonas aeruginosa ST395 950	6.55	66.1
Pseudomonas aeruginosa 19BR	6.74	-
Pseudomonas aeruginosa NCGM2.S1	6.76	66.1
Pseudomonas aeruginosa LESB58	6.6	66.3
Pseudomonas aeruginosa PAO1	6.26	66.6
Pseudomonas aeruginosa B136-33	6.42	66.4
Pseudomonas aeruginosa DK2	6.4	66.3
Pseudomonas aeruginosa M18	6.33	66.5
Pseudomonas aeruginosa PA7	6.59	66.4
Pseudomonas aeruginosa RP73	6.34	66.5
Pseudomonas aeruginosa UCBPP-PA14	6.54	66.3

Chart A.4.1. Genome features various *P. aeruginosa* strains (according to NCBI (2)). Current study strain highlighted.

Strain	Plasmids	Size, Mb	GC, %
Escherichia coli O104:H4 1218_280	3	5.41	50.6
Escherichia coli O104:H4 str. 2011C-3493	3	5.44	50.6
Escherichia coli O104:H4 str. 2009EL-2050	3	5.44	50.5
Escherichia coli O104:H4 str. 2009EL-2071	2	5.39	50.7
Escherichia coli 55989	-	5.15	50.7
Escherichia coli O26:H11 str. 11368	4	5.86	50.7
Escherichia coli O111:H- str. 11128	5	5.77	50.4
Escherichia coli O157:H7 str. EC4115	2	5.7	50.4
Escherichia coli UMNK88	5	5.67	50.7
Escherichia coli O157:H7 str. Sakai	2	5.59	50.4
Escherichia coli O103:H2 str. 12009	1	5.52	50.7
Escherichia coli O55:H7 str. CB9615	1	5.45	50.5
Escherichia coli O7:K1 str. CE10	4	5.38	50.6
Escherichia coli ETEC H10407	4	5.33	50.7
Escherichia coli UTI89	1	5.18	50.6
Escherichia coli O83:H1 str. NRG 857C	1	4.89	50.7
Escherichia coli ATCC 8739	-	4.75	50.9
Escherichia coli IAI1	-	4.7	50.8
Escherichia coli HS	-	4.64	50.8
Escherichia coli BL21(DE3)	-	4.56	50.8

Chart A.4.2. Genome features of different *E. coli* strains (according to NCBI (1)). Current study strain highlighted.

References

1. Agilent Technologies, Inc. (2007) Agilent DNA 7500 and DNA 12000 Kit Quick Start Guide. G2938-90025, Edition 04/2007: Germany
2. Agilent Technologies, Inc. (2009) Agilent High Sensitivity DNA Kit Guide. G2938-90321, Edition 5/09: Germany
3. Ahmed Sanaa A, Awosika J, Baldwin C, Bishop-Lilly K, Biswas B, Broomall S, Patrick C, Chertkov O, Chokoshvili O, Coyne S, Davenport K, Detter C, Dorman W, Erkkila T, Folster J, et al. (2012) Genomic Comparison of *Escherichia coli* O104:H4 Isolates from 2009 and 2011 Reveals Plasmid, and Prophage Heterogeneity, Including Shiga Toxin Encoding Phage stx2. Threat Characterization Consortium, PLoS One, 7(11): 48228
4. Aird D, et al. (2010) Analyzing and minimizing bias in Illumina sequencing libraries. Genome Biol., 11(Suppl. 1): 3
5. Aloush V, Navon-Venezia S, Seigman-Igra Y, Cabili S, Carmeli Y. (2006) Multidrug-resistant *Pseudomonas aeruginosa*: risk factors and clinical impact. Antimicrob. Agents Chemother., 50: 43–48
6. Arlt M, Cagla Ozdemir Shanda Birkeland A, Lyons R, Glover T, Wilson T. (2011) Comparison of Constitutional and Replication Stress-Induced Genome Structural Variation by SNP Array and Mate-Pair Sequencing. Genetics, 187(3): 675–683
7. Armitage J and Schmitt R. (1997) Bacterial chemotaxis: Rhodobacter sphaeroides and Sinorhizobium meliloti: D-variations on a theme? Microbiology, 143: 3671-3682
8. Baker M. (2012) *De novo* genome assembly: what every biologist should know. Nature methods, 9(4): 333-337
9. Ball C, et al. (2000) Integrating functional genomic information into the Saccharomyces genome database. Nucleic Acids Res., 28: 77-80
10. Bellais S, Aubert D, Naas T, Nordmann P. (2000) Molecular and biochemical heterogeneity of class B carbapenem-hydrolyzing β-lactamases in Chryseobacterium meningosepticum. Antimicrob. Agents Chemother., 44: 1878–1886
11. Bernardet J, Hugo C, Bruun B. (2006) The genera Chryseobacterium and Elizabethkingia. In The Prokaryotes. (Edited by: Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E.): New York, Springer, 7 (3rd edition): 638-676
12. Bloch K, Nadarajah R, Jacobs R. (1997) *Chryseobacterium meningosepticum*: an emerging pathogen among immunocompromised adults. Medicine, 76(1): 30–41

13. Boyle B, Fernandez L, Laroche J, Kukavica-Ibrulj I, Mendes C, Hancock R, Levesque R. (2012) Complete Genome Sequences of Three *Pseudomonas aeruginosa* Isolates with Phenotypes of Polymyxin B Adaptation and Inducible Resistance. J Bacteriol., 194(2): 529–530
14. Burgart L, Robinson R, Heller M, Wilke W, Iakoubova O, Cheville J. (1992) Multiplex polymerase chain reaction. Mod. Pathol., 5: 320–323
15. Caruccio N, Grunenwald H, Syed F. (2009) Nextera™ Technology for NGS DNA Library Preparation: Simultaneous Fragmentation and Tagging by In Vitro Transposition. Nextera Technology™. Vol. 16-3, EPICENTRE Biotechnologies, www.EpiBio.com
16. CBCB. (date unknown) Genome sequence assembly primer. 2005 – 2013, 3125 Biomolecular Sci., Bldg 296: University of Maryland, College Park, USA, [cited June 2013]. Available from: http://www.cbcb.umd.edu/research/assembly_primer.shtml
17. Ceyhan M, Yıldırım I, Tekelı A, Yurdakok M, Us E, Altun B, Kutluk T, Cengiz A, Gurbuz V, Barın C, Bagdat A, Cetınkaya D, Gur D, Tuncel O. (2008) A Chryseobacterium meningosepticum outbreak observed in 3 clusters involving both neonatal and non-neonatal pediatric patients. Am. J. Infect. Control, 36: 453-457
18. Ceyhan M. and Celik M. (2011) *Elizabethkingia meningosepticum* (*Chryseobacterium meningosepticum*). Infections in Children. Int. J. Pediatr., 215237
19. Chaisson M. and Pevzner P. (2008) Short read fragment assembly of bacterial genomes. Genome Res., 18: 324–330
20. Chamberlain J, Gibbs R, Ranier J, Nguyen P, Caskey C. (1988) Deletion screening of the duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Res., 16: 11141-11156
21. Costerton J, Stewart P, Greenberg E. (1999) Bacterial biofilms: a common cause of persistent infections. Science, 284: 1318-1322
22. Didelot X, Bowden R, Wilson D, Peto T, Crook D. (2012) Transforming clinical microbiology with bacterial genome sequencing. Nat. Rev. Genet., 13(9): 601-12
23. Dimatatac E, Alejandria M, Montalban C, Pineda C, Ang C, Delino R. (2003) Clinical outcomes and costs of care of antibiotic resistant *Pseudomonas aeruginosa* infections. Philipp. J. Microbiol. Infect. Dis., 32: 159–167
24. Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E. (2006) The Prokaryotes: Proteobacteria: Delta and Epsilon Subclasses. Deeply Rooting Bacteria. A Handbook on the Biology of Bacteria. Springer-Verlag, 7(3rd Edition): 653-665: New York, USA

25. Eyre D, et al. (2012) A pilot study of rapid benchtop sequencing of *Staphylococcus aureus* and *Clostridium difficile* for outbreak detection and surveillance. BMJ, Open 2: 001124
26. Fick R. (1993) *Pseudomonas aeruginosa*—the Microbial Hyena and Its Role in Disease: An Introduction. *Pseudomonas aeruginosa*. The Opportunist, 1-6
27. Fleischmann R, Adams M, White O, Clayton R, Kirkness E, Kerlavage A, Bult C, Tomb J, Dougherty B, Merrick J, McKenney K, Sutton G, FitzHugh W, Fields C, Gocayne J, Phillips C, et al. (1995) Whole genome random sequencing and assembly of *Haemophilus influenza.* Rd. Science, 269: 496–512
28. Frank C, Werber D, Cramer J, Askar M, Faber M, an der Heiden M, Bernard H, Fruth A, Prager R, Spode A, Wadl M, Zoufaly A, Jordan S, Kemper M, Follin P,Müller L, Krause G, et al. (2011) Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. N. Engl. J. Med. 10, 365(19): 1771-80
29. Girlich D, Naas T, Nordmann P. (2004) Biochemical characterization of the naturally occurring oxacillinase OXA-50 of *Pseudomonas aeruginosa*. Antimicrob. Agents Chemother., 48: 2043–2048
30. Grad Y, Lipsitch M, Feldgarden M, Arachchi H, Cerqueira G, Fitzgerald M, Godfrey P, Haas B, Murphy C, Russ C, Sykes S, Walker B, Wortman J,Young S, Zeng Q, Frim odtMoller J, et al. (2012) Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. Proceedings of the National Academy of Sciences, 109(8): 3065-3070
31. Hancock R. (1998) Resistance mechanisms in Pseudomonas aeruginosa and other nonfermentative gramnegative bacteria. Clin. Infect. Dis., 27: S93-S99
32. Hancock R. and Brinkman F. (2002) Function of pseudomonas porins in uptake and efflux. Annu. Rev. Microbiol., 56: 17–38
33. Hardalo C. and Edberg S. (1997) *Pseudomonas aeruginosa*: assessment of risk from drinking water. Crit. Rev. Microbiol., 23: 47-75
34. Harris S, et al. (2012) Whole-genome analysis of diverse *Chlamydia trachomatis* strains identifies phylogenetic relationships masked by current clinical typing. Nature Genet., 44: 413–419
35. Hauser A. and Ozer E. (2011) *Pseudomonas aeruginosa*. Nature Reviews Microbiology, 9: 3
36. Heiner C, Hunkapiller K, Chen S, Glass J, Chen E. (1998) Sequencing multimegabase-template DNA with BigDye terminator chemistry. Genome Res., 8: 557–561

37. Holmes B. and Owen R. (1981) Emendation of the genus Flavobacterium and the status of the genus. In Reichenbach H and Weeks O ed. The Flavobacterium-Cytophaga Group. Proceedings of the International Symposium on yellow-pigmented Gram-negative bacteria of the Flavobacterium-Cytophaga Group: Braunschweig, July 8 – 11, 1980. Weinheim, Verlag Chemie, 17-26
38. Hoque S, Graham J, Kaufmann M, Tabaqchali S. (2001) *Chryseobacterium (Flavobacterium) meningosepticum* outbreak associated with colonization of water taps in a neonatal intensive care unit. Journal of Hospital Infection, 47(3): 188–192
39. Illumina, Inc. (2012a) Nextera XT DNA Sample Preparation Guide. Illumina protocols. Part # 15031942 Rev. C
40. Illumina, Inc. (2012b) Data processing of Nextera Mate Pair reads on Illumina sequencing platforms. Technical Note: Sequencing. Pub. No. 770-2012-053
41. Illumina, Inc. (2013) Nextera Mate Pair Sample Preparation Guide. Illumina protocols. Part # 15035209 Rev. C
42. Jacobs A. and Hafizah C. (2011) Biofilm formation and adherence characteristics of an *Elizabethkingia meningoseptica* isolate from *Oreochromis mossambicus*. Annals. of Clin. Microbiol. and Antimicrob., 10: 16
43. Kämpfer P, Matthews H, Glaeser S, Martin K, Lodders N, Faye I. (2011) *Elizabethkingia anophelis* sp. nov., isolated from the midgut of the mosquito *Anopheles gambiae*. Int. J. Syst. Evol. Microbiol., 61(Pt 11): 2670-2675
44. Koser C, et al. (2012) Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N. Engl. J. Med., 366: 2267–2275
45. Kozarewa I, Ning Z, Quail M, Sanders M, Berriman M, Turner D. (2009) Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. Nat. Methods, 6: 291–295
46. Krieg N. (2006) Bergey's manual of systematic bacteriology. The genera Chryseobacterium and Elizabethkingia. 4: 202-209
47. Lagerstrom M, Parik J, Malmgren H, Stewart J, Pettersson U, Landegren U. (1991) Capture PCR: efficient amplification of DNA fragments adjacent to a known sequence in human and RAC DNA. PCR Methods Appl., 1: 111-119
48. Lederberg J, et al. (2000) *Pseudomonas*. Encyclopedia of Microbiology. 3(2nd Edition): 876-891, San Diego
49. Lieberman T, et al. (2011) Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. Nature Genet., 43: 1275–1280

50. Lister P, Wolter D, Hanson N. (2009) Antibacterial-Resistant *Pseudomonas aeruginosa*: Clinical Impact and Complex Regulation of Chromosomally Encoded Resistance Mechanisms. Clinical Microbiol. Reviews, 22(4) 582–610
51. Loman N, Constantinidou C, Chan J, Halachev M, Sergeant M, Penn C, Robinson E, Pallen M. (2012) High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. Nature Reviews, Microbiology, 10: 599-606
52. Lu C, et al. (2005) Elucidation of the small RNA component of the transcriptome. Science (New York, NY), 309: 1567–1569
53. Lu S, Zhang X, Zhu Y, Sik Kim K, Yang J, Jin Q. (2011) Complete Genome Sequence of the Neonatal-Meningitis-Associated *Escherichia coli* Strain CE10. J. bacterial., 7005
54. Malde K. (2008) The effect of sequence quality on sequence alignment. Bioinformatics (Oxford, England), 24: 897–900
55. Mardis E. (2008) The impact of next-generation sequencing technology on genetics. Trends Genet., 24: 133–141
56. Marine R, Polson Shawn W, Ravel J, Hatfull G, Russell D, Sullivan M, Syed F, Dumas M, Wommack E. (2011) Evaluation of a Transposase Protocol for Rapid Generation of Shotgun High-Throughput Sequencing Libraries from Nanogram Quantities of DNA. Appl. Environ. Microbiol., 77(22): 8071–8079
57. Mathee K, Giri N, Camilo V, Xiaoyun Q, Jody M, Koehrsen M, Rokas A, Yandava C, Engels R, Zeng E, Olavarietta R, Doud M, Smith R, Montgomery P, White J, Godfrey P, Lory S, et al. (2008) Dynamics of Pseudomonas aeruginosa genome evolution. PNAS, 105(8): 3100–3105
58. McAdam P, Holmes A, Templeton K, Fitzgerald J. (2011) Adaptive evolution of *Staphylococcus aureus* during chronic endobronchial infection of a cystic fibrosis patient. PLoS ONE, 6: 24301
59. McBride M, Weinberg R, Zusman D. (1989) ``Frizzy'' aggregation genes of the gliding bacterium *Myxococcus xanthus* show sequence similarities to the chemotaxis genes of enteric bacteria. Proc. Natl. Acad. Sci., 86: 424-428
60. Micek S, Lloyd A, Ritchie D, Reichley R, Fraser V, Kollef M. (2005) *Pseudomonas aeruginosa* bloodstream infection: importance of appropriate initial antimicrobial treatment. Antimicrob. Agents Chemother., 49: 1306–1311
61. Miquel S, Peyretaillade E, Claret L, de Valle´e A, Dossat C, et al. (2010) Complete Genome Sequence of Crohn's Disease-Associated Adherent-Invasive *E. coli* Strain LF82. PLoS ONE, 5(9): 12714

62. Miyoshi-Akiyama T, Kuwahara T, Tada T, Kitao T, Kirikae T (2011) Complete Genome Sequence of Highly Multidrug-Resistant *Pseudomonas aeruginosa* NCGM2.S1, a Representative Strain of a Cluster Endemic to Japan. J. Bacteriol., 193(24): 7010
63. Mizobuchi M and Frohman L. (1993) Rapid amplification of genomic DNA ends. BioTechniques, 15: 214
64. Murphy S, Cheville J, Zarei S, Johnson S, Sikkink R, Kosari F, Feldman A, Eckloff B, Karnes J, Vasmatzis G. (2012) Mate Pair Sequencing of Whole-Genome-Amplified DNA Following Laser Capture Microdissection of Prostate Cancer. DNA Res., 19(5): 395–406
65. Mutreja A, et al. (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. Nature, 477: 462–465
66. Nagarajan N, Cook C, Di Bonaventura M, Ge H, Richards A, Bishop-Lilly K, DeSalle R, Read T, Pop M. (2010) Finishing genomes with limited resources: lessons from an ensemble of microbial genomes. BMC Genomics, 11: 242
67. NCBI (1) genome browser. (date unknown) Genomes & Maps: Genome: *E. coli.* National Center for Biotechnology Information, U.S. National Library of Medicine: Rockville Pike, Bethesda MD, 20894, USA, [cited June 2013]. Available from: http://www.ncbi.nlm.nih.gov/genome/167
68. NCBI (2) genome browser. (date unknown) Genomes & Maps: Genome: *P. aeruginosa* National Center for Biotechnology Information, U.S. National Library of Medicine: Rockville Pike, Bethesda MD, 20894, USA, [cited June 2013]. Available from: http://www.ncbi.nlm.nih.gov/genome/?term=pseudomonas+aeruginosa
69. Nielsen R, Paul J, Albrechtsen A, Song Y. (2011) Genotype and SNP calling from next-generation sequencing data. Nature Rev. Genet., 12: 443–451
70. Nieuwerburgh F, Thompson R, Ledesma J, Deforce D, Gaasterland T, Ordoukhanian P, Head S. (2011) Illumina mate-paired DNA sequencing-library preparation using Cre-Lox recombination. Nucl. Acids Research, 40(3): 24
71. Nikaido H. (1998) Antibiotic resistance caused by gram-negative multidrug efflux pumps. Clin. Infect. Dis., 27: S32-S41
72. Pollack M. (1995) *Pseudomonas aeruginosa*. In G. L. Mandell, R. Dolan and J. E. Bennett (ed.), Principles and practices of infectious diseases: Churchill Livingstone, New York, 1820–2003
73. Poole K. (2005) Aminoglycoside resistance in *Pseudomonas aeruginosa*. Antimicrob. Agents Chemother., 49: 479–487

74. Pop M. and Salzberg S. (2008) Bioinformatics challenges of new sequencing technology. Trends Genet., 24: 142–149
75. Rasko D, et al. (2011) Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. N. Engl. J. Med., 365: 709–717
76. Relman D. (2011) Microbial genomics and infectious diseases. N. Engl. J. Med., 365: 347–357
77. Rogers Y, Munk C, Meincke L, Han C. (2005) Closing bacterial genomic sequence gaps with adaptor-PCR. BioTechniques, 39: 31-34
78. Rohde H, Qin J, Cui Y, Li D, Loman N, Moritz H, Wentong C, Fei P, Yangqing P, Li J, Feng X, Li S, Li Y, et al. and the *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium. (2011) Open-Source Genomic Analysis of Shiga-Toxin–Producing *E. coli* O104:H4. N. Engl. J. Med., 365: 718-724
79. Sanger F, Nicklen S, Coulson A. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci., 74: 5463–5467
80. Sarkar G, Turner R, Bolander M. (1993) Restriction-site PCR: a direct method of unknown sequence retrieval adjacent to a known locus by using universal primers. PCR Methods Appl., 2: 318-322
81. Siebert P, Chenchik A, Kellogg D, Lukyanov K, Lukyanov S. (1995) An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res., 23: 1087-1088
82. Soon W, Hariharan M, Snyder M. (2013) High-throughput sequencing for biology and medicine. Molecular Systems Biology, 9: 640
83. Spencer R. (1996) Predominant pathogens found in the European prevalence of infection in intensive care study. Eur. J. Clin. Microbiol. Infect. Dis., 15: 281–285
84. Steinberg J. and Burd E. (2010) Other gram-negative and gram-variable bacilli. In: Mandell G, Bennett J, Dolin R. edition. Principles and Practice of Infectious Diseases: Philadelphia, Churchill Livingstone Elsevier, 3015-3033
85. Steinberg J. (2000) Other gram-negative bacilli. In: Mandel G, Bennett J, Dolin R. edition. Principles and Practice of Infectious Diseases: Churchill Livingstone, Edinburgh, UK, 2459–2474
86. Stock J. and Surette M. (1996) Chemotaxis in *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology (ed. Neidhardt F.), 1103-1129: ASM, Washington, DC
87. Stover C, Pham X, Erwin A, Mizoguchi S, Warrener P, Hickey M, Brinkman F, Hufnagle W, Kowalik D, Lagrou M, Garber R, Goltry L, Tolentino E, Westbrock-

Wadman S, Olson M, et al. (2000) Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. Nature, 406: 959-964

88. Suma A, Kumar N, Baddam R, Arif H, Nishant N, Savita J, Niyaz A. (2011) Genome of Multidrug-Resistant Uropathogenic *Escherichia coli* Strain NA114 from India. J. Bacteriol., 4272–4273
89. Tettelin H, Radune D, Kasif S, Khouri H, Salzberg S. (1999) Optimized Multiplex PCR: efficiently closing a whole-genome shotgun sequencing project**.** Genomics, 62(3): 500-507
90. Triglia T, Peterson M, Kemp D. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences. Nucleic Acids Res., 16: 8186-8186
91. Walenz B, Sutton G, Miller J. (2011) Pair classification within Illumina mate pair data. DeNovo Classification Poster from CSH Genome Informatics. The J. Craig Venter Institute, Rockville, USA
92. Westbrock-Wadman S, et al. (1999) Characterization of a Pseudomonas aeruginosa efflux pump contributing to aminoglycoside impermeability. Antimicrob. Agents Chemother., 43: 2975-2983
93. Winstanley C, Langille M, Fothergill J, Kukavica-Ibrulj I, Paradis-Bleau C, Sanschagrin F, Thomson N, Winsor G, Quail M, Lennard N, Bignell A, Clarke L, Seeger K, Levesque R, et al. (2009) Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool Epidemic Strain of *Pseudomonas aeruginosa.* Genome Res., 19(1): 12–23
94. Yuan J, Liu W, Sun M, Song S, Cai J, Hu S. (2011) Complete Genome Sequence of the Pathogenic Bacterium *Riemerella anatipestifer* Strain RA-GD. J Bacteriol., 193(11): 2896–2897
95. Zerbino D. and Birney E. (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res., 18: 821–829