UNIVERSITATEA „ALEXANDRU IOAN CUZA”
DIN IAŞI

JOURNAL OF EXPERIMENTAL AND MOLECULAR BIOLOGY

TOME XXII, Number 1 **2021**

Editura Universităţii „ALEXANDRU IOAN CUZA” din Iaşi

EDITORIAL OFFICE

Universitatea „Alexandru Ioan Cuza" din Iași, Facultatea de BIOLOGIE
Laboratorul de Biochimie și Biologie Moleculară
Bulevardul Carol I, Nr. 20A, 700506, Iași, România
www.jemb.bio.uaic.ro / gbmpapers@yahoo.com

CONTENT

DOI:10.47743/jemb-2021-54

EPIDEMIOLOGICAL CHARACTERISTICS OF CHILDREN'S POISONING WITH HOUSEHOLD SUBSTANCES - RETROSPECTIVE STUDY

NICOLAI NISTOR[1,2], DRAGOȘ CRAUCIUC[1,3], TATIANA IOV[3], EDUARD CRAUCIUC[1,4*], DIANA POPOVICI[1,4], CRISTINA JITĂREANU[2], LOREDANA LILIANA HURJUI[1,5], ANDRA IRINA BULGARU-ILIESCU[1]

Received: 2nd of February 2021 / Revised: 20th of June 2021
Accepted: 29th of June 2021 / Published: 22nd of March 2022

Keywords: household substances, intoxication, children.

Abstract. Poisoning with household substances is responsible for a significant morbidity in children. The objective of this study is to specify the epidemiological characteristics of accidental and voluntary poisoning with these products at pediatric age. The study was observational, retrospective and included acute intoxication by ingestion of household substances in children aged 0-18 years hospitalized in The Regional Toxicology Center of the Emergency Hospital for Children "St. Maria "Iasi. The study group included 230 children admitted in the last 6 years with this diagnosis. Data were obtained from patient observation sheets, centralized and subsequently processed statistically. Knowing the epidemiological aspects of these poisonings can contribute to the elaboration and application of appropriate prevention strategies

INTRODUCTION

Acute poisoning in children is a worldwide public health problem, but it can be avoided. The major responsibility for preventing this lies with the family (Iov, T., et al, 2019). Acute intoxication in children is very different from that of the adults in the ways it happens, the incriminating toxins, the methods of approach and especially the consequences. The child with acute intoxication is never a miniature adult; they react totally differently at the impact with different toxic substances. The most important difference with regard to child versus adult poisoning is the causative agent. While in adults, drug-related toxins (antidepressants, sedatives, anxiolytics) are most often incriminated, in children non-drug toxins, including household substances, are prevalent (Qazy, M. & Saqib, N., 2018). Epidemiological studies on acute intoxication in children have found that the most frequent risk factors are the small age, poor level of family education and low socio-economic status. In the pediatric age there is a bimodal distribution with a peak frequency in the young child (accidental intoxication especially in boys) and the second peak in adolescence (voluntary intoxication predominantly in girls (Fan, A.Y., et al, 2013).

The products used in the household are often within the reach of children in the age group 1-5, the age at which they explore the environment. The main toxic substances in this category are in the kitchen, bathroom and toilet. Household products are usually beautifully colored liquids that tempt children. At this age they are hyperactive and express their exploratory behavior, but they are too young to realize which substances are dangerous (Calello, D.P. & Henretig, F.M., 2014).

In the last years there has been a permanent change in the spectrum of toxic substances incriminated in accidental or voluntary intoxication in children, which requires the permanent information of the doctors on its epidemiological aspects. Although in most cases acute poisoning with these substances is benign, in some cases, especially when ingesting corrosive substances containing strong acids or bases, severe chemical burns can occur in the digestive tract starting from the mouth to the stomach. Corrosives are included in the composition of solutions for cleaning household objects, strippers and decanters. In addition to accidental poisoning at an early age, these solutions are also involved in adolescent voluntary intoxication, sometimes with the purpose of commiting suicide (Faz, A.A., et al, 2017).

The identification and documentation of the epidemiological aspects in these intoxications are of particular importance for the application of appropriate preventive measures (Prasadi, G.A.M., et al, 2018).

The objective of this study was to determine the risk factors in accidental or voluntary intoxications in children.

The high frequency of acute intoxication in the age group 1-4 can be explained by the following aspects:

- after one year of age children begin to walk, explore the environment, start to open the cabinets and examine their contents;
- the tendency of the child is to introduce the new objects he examines into the mouth cavity;
- the skill level increases, and the child begins to easily unscrew the caps of medicine bottles or other containers;

DOI:10.47743/jemb-2021-54

- parents' negligence in storing and using the toxic substances; they are either kept in containers similar to those containing food or juices, or they are not stored properly, or when used they become accessible to the child.
- Hydroelectrolytic disorders are often encountered in the acute intoxication of the child, some of which are due to the toxicity. Knowing these hydroelectrolytic disorders and their proper treatment helps reduce mortality in children with acute intoxication.

MATERIAL AND METHOD

We conducted an observational, retrospective study on acute intoxication with household substances in children aged 0-18, hospitalized in the Regional Toxicology Center of the Emergency Clinical Hospital for Children "Sf. Maria" Iasi. The study group included 230 children admitted in the period 1.01.2014 - 31.12.2019 with the diagnosis of poisoning by ingesting one of the products in this class. Children with inhalation poisoning as well as those who could not specify the toxic substance ingested were excluded from the batch. The data obtained from the patient observation sheets were centralized in a SPSS 18.0 database and subsequently processed, taking into account the relative risk (RR) and the confidence interval 95% (CI 95%).
Data obtained from patients' files was gathered and processed in a SPSS 18.0 database. Discrete variables were expressed as number and proportion. The comparison of the groups for the categorical variables was done by the Chi square test ($\chi 2$), the significance threshold being $P=1$. We evaluated the relative risk (RR) and the 95% confidence interval (CI 95%), following statistical interpretation: $RR>1$, the variable represents a risk factor; $RR=1$, the variable represents an indifferent factor. The statistical signification was accepted at a value of $P<.05$. The study was approved by the Ethics Committee of the Hospital.

RESULTS AND DISCUSSIONS

During the study period, a total of 2915 children with acute intoxication were admitted, of which 230 (7.8%) were intoxicated with household substances. Over the course of the 6 years, the percentage of poisonings with household substances from the total poisoning ranged from 6.8% in 2015 to 8.6% in 2018 (Table I).

Table I. Annual distribution of the number of intoxications with household substances from the total number of intoxications

Year	2014	2015	2016	2017	2018	2019
Total number of intoxications	529	528	510	480	464	404
Number of cases of intoxication with household substances	42 (7.9%)	36 (6.8%)	43 (8.4%)	36 (7.5%)	40 (8.6%)	33 (8.2%)

The frequency of intoxication with household substances varies from country to country and remains difficult to assess because in many cases it is a benign intoxication for which patients do not show up for a medical consult. Over the 6 years, the average percentage of the cases involving household substances was 7.8% in the context of the decrease in the last years of the total number of poisonings. The frequency is slightly above the 7.6% reported by Lee et al., 2019 and lower but statistically insignificant compared to two other studies reporting 10.2% by Huynh et al., 2018 and 10% Devaranavadagi et al., 2017 respectively. Another study based on the prospective collection of telephone calls for various poisonings found that 10% of them involved household substances (Williams, H., et al., 2012).
Intoxications with more common household substances were accidental (168/73%) compared to voluntary (62/27%). Table II presents the demographic aspects of accidental poisoning with these substances in the study group.

DOI:10.47743/jemb-2021-54

Table II. Distribution of accidental intoxication by age, gender and environment

Age	< 1 year old	1 -2 years old	2-3 years old	3-4 years old	4-5 years old	>5 years old	Male	Female	From the urban area	From the rural area
Number of cases	5	42	61	28	14	18	104	64	52	116
Percentage	3%	25%	36%	17%	8.3%	10.7%	62.5%	37.5%	31%	69%

Most cases of accidental poisoning with household substances have been found at the age of 2-3 years old (36%). Other studies report that at this age intoxications with such substances were more common, as well (Manzar, N., et al., 2010). There are also studies that found the highest frequency (39%) at the age of 1-2 (Vilaçaa, L., et al., 2020). In the study group, these intoxications were more frequent in female children and in those living in rural areas.
Voluntary poisoning was registered in 62 patients. The average age at which these poisonings were reported was 15 ± 2.05 years old.

Table III. Distribution of voluntary intoxications by gender and environment.

	Male	Female	From the urban area	From the rural area
Number of cases	14	48	28	34
Percentage	21.5%	78.5%	45.2%	54.8%

These poisonings, sometimes use for committing suicide, were more common in girls (78.5%), in line with the current trend of voluntary intoxication in adolescents, which is 73.6% (Sheridan, D., et al., 2019). There was a slight predominance in children living in rural areas.
The main household substances involved in acute intoxication in children in this group are detergents, liquid soap, shower gel, body deodorizers, nail polish and solvents, strippers and detergents, sewage products, household pesticides and ethylene glycol (table IV).
The ingested household products can be corrosive, irritating or just foaming. They have four forms: liquids, gels, solids and powders (Fieux, F., et al., 2013). In the accidental poisonings of the study group, detergents (58 cases) were most frequently involved. Most detergents contain anionic surfactants and various additives. Their toxic action is achieved by more or less severe digestive irritation depending on the amount ingested. In recent years, they are commonly used packed in cushions containing 32-50 ml of concentrated liquid detergent, wrapped in a water-soluble membrane, providing a correct dose for washing machines or dishes and a reduction in packaging size (Settimi L, et al., 2018). Second as frequency, decapitators and decanters (22 cases) and household pesticides (20 cases) were involved. We noticed smaller percentages of poisoning with sewage products, liquid soap, shower gel, ethylene glycol, body deodorizers, nail polish and solvents.

DOI:10.47743/jemb-2021-54

Table IV. Involving household substances in accidental and voluntary intoxication in children

Type of intoxication	**Accidental poisoning**	**Voluntary poisoning**	**Chi²**	**p**	**RR** A-accidental V-voluntary	**CI 95%**
Toxic substances involved	168	62				
Detergents	58 (34.5%)	16 (25.8%)	1.20	0.272	1.34_A	0.84-2.14
Liquid soap	14 (8.3%)	0 (0%)	5.48	0.019	8.33_A	4.15-12.5
Shower gel	12 (7.1%)	0 (0%)	4.65	0.031	7.14_A	3.25-11.0
Body deodorizers	10 (6.0%)	0 (0%)	3.84	0.050	5.95_A	2.37-9.53
Nail polish and solvents	5 (3.0%)	0 (0%)	1.88	0.171	2.98_A	0.41-5.55
Strippers and descaling agents	22 (13.1%)	14 (22.6%)	3.07	0.800	1.72_V	0.94-3.15
Products for draining the sewerage	15 (8.9%)	10 (16.1%)	2.41	0.120	1.81_V	0.86-3.81
Household pesticides	20 (11.9%)	18 (29.0%)	9.59	0.002	2.44_V	1.38-4.30
Ethylene glycol	12 (7.1%)	4 (6.5%)	0.06	0.804	1.15_A	0.38-3.43

Accidental poisoning with liquid soap (RR=8.33; CI95%: 4.15-12.5), shower gel (RR=7.14; CI 95%: 3.25-11.0) and body deodorizers (RR=5.95; CI95%: 2.37-9.53) presents a significantly increased risk in the young child, while poisoning with nail polish and solvent (RR = 2.98; 95% CI 0.41-5.55) presents a risk about 3 times higher, but the result cannot be extrapolated to the general pediatric population. In the voluntary intoxication, only five types of household substances were involved, which in descending order were household pesticides (18 cases), detergents (16 cases), strippers and decanters (14 cases), sewage products (10 cases) and ethylene glycol (4 cases). In these voluntary poisonings, a large number of corrosive substances were involved and they can cause severe damage to the digestive tract. The severity of these lesions depends on the chemical characteristics of the substance involved (acid or base), its dilution, the amount ingested and its presentation form (solution, tablets, granules or flakes) (Nițescu, V., 2015). In all cases of ingestion of such substances, it is essential to call an antitoxic center, as the composition and concentration of these products are sometimes not clearly specified on labels and they may be made of several corrosive agents or contain substances with systemic toxicity (Fieux, F., et al., 2013). The cases of poisoning by ingestion of corrosive substances in the study group were treated in the acute stage but we do not know the number of those with late complications, as they were followed in the pediatric gastroenterology section. The most common late complication is post-caustic esophageal stenosis, which occurs 3 weeks after ingestion. A study that was conducted on a large group of children with ingestion of caustic substances found the presence of esophageal stenosis in 13.5% of cases (Karaman, I., et al., 2015). The involvement of ethylene glycol in 12 cases of accidental poisoning and 4 cases of voluntary poisoning should also be mentioned. It is part of several chemicals available to children in the home environment, the most popular being antifreeze and brake fluid. Although the cases in our group have evolved favorably, sometimes these intoxications have severe or even

DOI:10.47743/jemb-2021-54

fatal prognosis. The toxic dose in children is 0.7 ml*kg^{-1} and the lethal dose is 1.5 ml*kg^{-1} (Furnica, C., et al., 2017).

Although it includes a significant number of patients, this study has some limitations. One of them is the retrospective character, the analyzed data being obtained from the observation sheets of the admitted patients, a method that is far from perfect. Another limitation is the inclusion in the study only of the children admitted with intoxication with household substances, but there were patients with benign forms, who were taken care of at home or in the emergency room, without being admitted.

CONCLUSIONS

Acute intoxications with household substances accounted for 6.8% of all cases of intoxication in children admitted during the respective period, 73% being accidental. The products involved were varied, most commonly being detergents, but corrosive substances that cause severe digestive damage were also found in a significant percentage. Knowing the epidemiological aspects of these poisonings can contribute to the elaboration and application of appropriate prevention strategies

REFERENCES

Iov, T., Knieling, A., Damian, S.I., et al. (2019). Domestic Poisoning in the Pediatric Age Group. Rev Chim (Bucharest), 70 (12): 4163-4168.

Qazi, M., Saqib, N. (2018). Clinical Profile and Outcome of Children Presenting with Poisoning or Intoxication: A Hospital-Based Study. Int J Contemp Pediatr. **5**(5): 1844-1850.

Fan, A.Y., Che Liu, A.H., Pan, B., et al. (2013). Investigating Childhood and Adolescence Poisoning Exposures in New Zealand Reported to the National Poisons Centre during 2000-2009. Asia Pacific J Med Tox, 52-57.

Calello, D.P., Henretig, F.M. , et, al. (2014). Pediatric Toxicology Specialized Approach to the Poisoned Child. Emerg Med Clin N Am. 32: 29–52.

Faz, A.A., Arsan, F., Peyvandi, M., et al. (2017). Epidemiologic Features and Outcomes of Caustic Ingestions; A 10-Year Cross-Sectional Study. Emergency, 5(1): e56.

Prasadi, G.A.M., Mohamed, F., Senarathna, L., et al. (2018). Paediatric Poisoning In Rural Sri Lanka: An Epidemiological Study. BMC Public Health. 18: 1349.

Lee, J., Fan, N.C., Yao, T.C., et al. (2019). Clinical Spectrum Of Acute Poisoning Inchildren Admitted To The Pediatric Emergency Department. Ped Neonat. 60: 59-67.

Huynh, A., Cairns, R., Brown, J.A., et al. (2018). Patterns of Poisoning Exposure at Different Ages: The 2015 Annual Report of the Australian Poisons Information Centres. Med J Aust 2018; 209(2): 74-79. || doi: 10.5694/mja17.01063.

Devaranavadagi, R.A., Patel, S., Shankar, P. (2017). A Study On Profile Of Poisoning In Pediatric Population. Int J Contemp Pediatr. 4(3): 810-815.

Williams, H., Moyns, E., Bateman, N., et al. (2012). Hazard of Household Cleaning Products: A Study Undertaken by the UK National Poisons Information Service. Clin Toxic, 50(8): 770-775.

Manzar, N., Ali Saad, S.M., Syeda, S.F., et al. (2010). The Study of Etiological and Demographic Characteristics of Acute Household Accidental Poisoning in Children - A Consecutive Case Series Study From Pakistan. BMC Ped. 10:28.

Vilaçaa, L., Volpea, F.M., Ladeiraa, R.M. (2020). Accidental Poisoning in Children and Adolescents Admitted to a Referral Toxicology Department of a Brazilian Emergency Hospital. Rev Paul Pediatr. 38: e2018096.

Sheridan, D., Hendrickson, R., Beauchamp, G., et al. (2019). Adolescent Intentional Abuse Ingestions: Overall 10-Year Trends and Regional Variation. Ped Emerg Care. 35(3): 176–179.

Fieux, F., Le Gall, C., Chirica, M., et al. (2013). Ingestion de produits ménagers (caustiques essentiellement) : évaluation de la gravité et du pronostic. Urgences. 1-20.

Settimi L, Giordano F, Lauria L,, et al. (2018). Surveillance of Paediatric Exposures to Liquid Laundry Detergent Pods in Italy. Inj Prev. 24: 5–11. doi:10.1136/injuryprev-2016-042263.

Nițescu, V. (2015). Intoxicațiile Acute cu Substanțe Caustice. În Ulmeanu C, Nițescu V. Intoxicațiile Acute la Copil și Adolescent. Ed. Tridona, 305-312.

Karaman, I., Koç, O., Karaman, A., et al. (2015). Evaluation of 968 Children with Corrosive Substance Ingestion. Indian J Crit Care Med. 19: 714-8.

DOI:10.47743/jemb-2021-54

Furnica, C., Knieling, A., Damian, S.I., et al. (2017). Fatal Ethylene Glycol Intoxication Secondary to Accidental Ingestion. Rev Chim (Bucharest), 68(7): 1591-1594.

1. University of Medicine and Pharmacy "Grigore T. Popa" Iași, România
2. Pediatric Department "St. Mary" Children Emergency Hospital, Iași, România
3. The Institute of Forensic Medicine, Iași, România
4. "Elena Doamna" Iași Clinical Hospital
5. Department of Morpho-Functional Sciences II, Discipline of Physiology

* crauciuc@yahoo.com

DOI:10.47743/jemb-2021-55

COLLATERAL PROSPECTS OF POLYHERBAL FORMULATION AS ANTIDIABETIC AND APPRAISING ITS TOXICITY IN EARLY DEVELOPMENT OF ZEBRAFISH (*DANIO RERIO*) LARVAE

AVILA JERLEY[1], ANITA MARGRET[2*], PAVITHRA LAKSHMI[1]

Received: 26th of June 2021 / Revised: 10th of July 2021
Accepted: 31st of August 2021 / Published: 22nd of March 2022

Keywords: Anti-diabetic activity, Polyherbal formulation, phytochemical, FASAME, zebrafish larvae.

Abstract. To formulate a polyherbal formulation, evaluate their toxicity and antidiabetic potential level using zebrafish model.

INTRODUCTION

Diabetes mellitus is a chronic disease that last long with several disorder associated both metabolically and physically.[1] The control of diabetes mellitus is challenging and it is a universal problem. The successful natural approach of treating these diseases has paved the new way in drug discovery.[2] Although there are many drugs which are consumed orally or injected intravenously to control the blood sugar level, they may develop some side effects and are also of high cost. This condition leads to exploration of the traditional herbal plants which can have complete cure without any side effects. The Ayurvedic medicine has a primeval knowledge and considered to be the ancient remedial sciences that has persisted over many generations in many countries as traditional medicine. In India it was prevailed thousands of years ago, and it is acknowledged as the "Mother of All Healing".[3] In this system of medicine, nearly 600 different plants have capacity to overcome diabetes and it is scientifically reported in some of the Indian books like Charak, Samhita, Mahdhav Nidan and Astang Sanghra.[4] There are more than 12,000 plants that have shown good have ethnopharmacological effect against diabetes, but most of them lack in scientific authentication.[5]Recently certain therapies are using oral hypoglycaemic tablets which does not restores the normal glucose level and also the have side effects.[6] This is considered as an immense challenge which needs to resolved by exploring the efficiency of herbs which divulge conventional antidiabetic activity . The World Health Organization also supports and recommends several medicinal plants and investigating its potent in combination with modern medicines.[7] Herbs and phytochemicals production plays a main role in the detection of novel therapeutic compounds and also, they possess antioxidants, hypoglycemic, and antihyperlipidemic properties. Sharangdhar Samhita, an ayurvedic literature from 1300AD has noted the importance of polyherbalism.[8] Polyherbal formulations (PHFs) boost the therapeutic activity and condense the concentrations of single herbs and decreases the adverse effects. The PHF has enhanced and multi-targeted therapeutic properties than the single herb. The idea of polyherbal is different and it is evident that many herbs or medicinal plants have various bioactive compounds with different mode of mechanism. So, this can give a good synergic response and can act as effective drug to treat several disorders. The efficacy of new drugs is generally authenticated and promoted by clinical trials using animal models. The zebrafish model has become prevalent to carry out drug efficacy studies due to its desirable morphological characters like small, transparent body and large clutch size. Further it exhibits a greater genomic and physiological resemblances to humans .Drug monitoring and accessing its effects can be performed extensively due to its transparent body features which can optimize optimal in vivo chemical screening and imaging. Hence these features enhance zebrafish as an optimal model to study developmental, physiological, and pathological processes [9]. Further various human diseases can be mimicked to understand their complications and mechanisms [10]. The present study is modulated to formulate a polyherbal formulation for antidiabetic activity and evaluate its toxicity using zebra fish as animal model.

MATERIALS AND METHODS

MATERIALS AND METHODS

Collection of plant materials and Preparation of polyherbal formulation - FASAME

The plants like *Ficus religiosa* (leaves), *Allium sativum* (Bulb), *Senna auriculata* (Flower) *Andrographis paniculata* (leaves), *Momordica charantia* (Fruit), *Eugenia jambolana* (leaves) were collected in January to February from various places in Tiruchirappalli, Tamil Nadu (Fig. 1).

DOI:10.47743/jemb-2021-55

Figure 1. Collection of Herbal plants A. *Ficus religiosa (leaves), B. Allium sativum (Bulb), C. Momordica charantia (Fruit) D. Senna auriculata (Flower) E. Eugenia jambolana (leaves) F. Andrographis paniculata (leaves)*

These plants were identified based on their anatomical and morphological features at the Rapinat Herbarium and Centre for Molecular Systematic, St. Josephs College (Autonomous), Tiruchirappalli, Tamilnadu, India where, the voucher specimen number (AAJ 006) was assigned and deposited. The collected materials were washed thoroughly in running tap water to remove soil particles and other debris. The leaves, fruit, bulb and flowers of this specific plants were shade dried separately and ground to a fine powder using electric blender. The powdered materials were stored in an air tight container for further investigation.

Phytochemical extraction

Different ratio of powdered samples of *Ficus religiosa* – 15 g, *Allium sativum* – 5 g, *Senna auriculata* – 10 g, *Andrographis paniculata* – 10 g, *Momordica charantia* – 5 g and *Eugenia jambolana* – 5 g are used for extraction by soxhlet apparatus (Accumax ,India) at boiling temperature (78.2 ° C) and the solvent used for extraction was hydro alcohol which constituted of analytical grade ethanol, methanol and distilled water at 7 : 3. ratio for 1 hr (Fig.2).

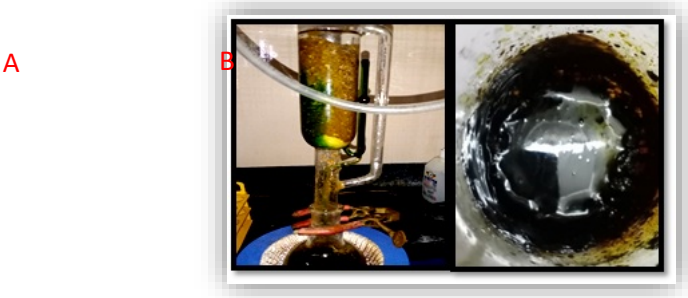

Figure 2. Extraction and formulation of FASAME - a polyherbal mélange
A-Soxhlet set up loaded with poly herbal hydroalcoholic suspension B- Condensed extract

The crude extract obtained were assessed for the qualitative phytochemical characterization for the identification of the various classes of active chemical constituents like resins, carboxylic acid, flavonoids, tannins, steroids, carbohydrates, glycosidase saponification, proteins, phenol, saponins, and gums using standard prescribed methods [11-13]. The positive tests were noted as weak (+), moderate (++), strong (+++) and absent (-).

Detection of Antidiabetic activity by β-galactosidase Assay

For the inhibition of β-galactosidase activity, a total of 0.5 mL of the different concentration of FASAME sample was pre-incubated with β-galactosidase in Na-acetate buffer at room temperature for 20 min. Then 0.5 mL of substrate mixture (8.3 mM ortho nitrophenyl β-D-galactopyranoside, 1 mM $MgCl_2$, and 0.1 M β-mercaptoethanol in 0.1 M Na-

DOI:10.47743/jemb-2021-55

phosphate buffer, pH 7.0) was added to the sample mixture. After the incubation at 30°C for 20 min, the reaction was terminated with 0.5 mL of 0.5 m Na_2CO_3 buffer. Release of o-nitrophenol was recorded at 420 nm using a microplate reader (Multiskan™ FC Microplate Photometer, Thermo scientific, USA). Each measurement was performed in triplicates using 96 well plate. Inhibition of enzyme activity was determined by using the following formula.

$$\text{Percentage of inhibition (\%)} = \frac{\text{Absorbance of control- Absorbance of test sample}}{\text{Absorbance of control}} \times 100 \quad (1)$$

Zebrafish Maintenance and Embryo Collection

Adult wild type zebrafish (Danio rerio) at 3-months of age of mixed sex were purchased from a local supplier (NSK Aquarium, Kolathur, Chennai,Tamilnadu,India).The fishes were maintained in closed ,9-L glass tanks (20 fish per tank) with charcoal-filtered tap water that endured the following physical parameters of 28 (± 0.5) ˚C temperature , pH 6.8–7.5 ,dissolved oxygen > 4 mg/L, conductivity 500–650 µS cm^{-1} ,total dissolved salt (TDS) 250 to 325 ppm) and salinity of 800–1200 µS/cm. An alternate 14/10 h light/dark cycle conditions were provided and Brine shrimp was feed twice daily at 9:00hrs and 16:00hrs. At night before fertilization, adult male and female zebrafish in 2:1 ratio were placed on opposite sides of a divider in a breeding tank. The next day morning, the zebrafish laid their eggs by natural mating soon after the exposure of first light. Embryos were collected within 30 min after spawning and rinsed with fresh water three times. The clean embryos were moved to tanks with the prepared embryo medium (14.61 g NaCl, 0.63 g KCl, 2.43 g mM $CaCl_2$ $2H_2O$, and 1.99 g $MgSO_4$, in 1 L of deionized water; stored at room temperature (~20–25° C) ;Diluted to 1X in deionized water prior to use) and cultured at 28°C for the subsequent experiments.

Zebrafish Embryo Toxicity

The test was carried out according to Organization for Economic Co-operation and Development (OECD) Test No. 236, 2013 (TIB Reg.No.219 /Res.25-09/09/2020). The median inhibitory concentration (IC_{50}) of the formulation was determined for 24hrs prior to the embryo toxicity studies .After elucidating an optimal threshold value, the healthy larvae were randomly divided into 24-well plates (5/well) and exposed to various concentrations of polyherbal formulation (0, 50, 100, 250 and 500 µg/mL) in 5 mL of fish water, in triplicate. The exposure lasted for 1 h in an environment at 28 ± 0.5°C under the same light/dark cycle (Fig. 3). After 1 h the solutions were replaced and dead embryos were discarded. The mortality rates were recorded, and examined using a stereomicroscope.

Morphology Score

Morphology scores were determined at 20 h post-treatment. Nine endpoints, including body shape, somite, notochord, tail, fins, heart, face, brain, and pharyngeal arches/jaws, were examined to evaluate the phenotypes of the zebrafish, and eight larval specimens per group were used for scoring. Subsequently, the larvae were anesthetized with 0.25 mg/mL tricaine and were observed and photographed using an inverted microscope.

Behavioral Analysis

The locomotor response of the larvae to physiological stress was determined after exposing them to the formulation for 2hrs. Further the larvae (sets of 5) were transferred to 24-well plates with 0.5 mL of fish water and incubated for 30 min at 28°C. The behavioral study was observed for a duration of 10 mins. The swimming pattern was considered as the foremost parameter for the survival of larvae.

Statistical analyses

Data were analysed using SPSS software (Version 24) The results were presented as mean _ standard deviation, while the results with a p-value below 0.05 were considered significantly different.

RESULTS AND DISCUSSIONS

In this work we focused to develop a polyherbal formulation for the treatment of diabetes. And also to prevent the side effects that is been caused by use of allopathy medicines and to reduce the financial crisis of the diabetic person [14]. Therefore, we tested the antidiabetic effect of FASAME, a polyherbal formulation by β-galactosidase assay and their toxicity using zebra fish model. The FASAME is a mixture of phytochemicals that was extracted from several plants like *Ficus religiosa, Allium sativum, Senna auriculata, Andrographis paniculata, Momordica charantia, Eugenia jambolana*. The soxhlet apparatus are used to fraction this poly herbal mixtures.The phytochemical studies revealed the presence of flavonoids, steroids, terpenoid, saponins, resins, carbohydrates, proteins and essential oil in the hydroalcoholic extract of FASAME polyherbal formulation. (Supplementary Table 1). The inhibitory concentration of FASAME was calculated for toxcitity testing in zebra fish (Supplementary Table 2).

DOI:10.47743/jemb-2021-55

These plants have antidiabetic property and they exhibit various mechanisms. Some plants can synthesis insulin secretion from beta cells of islets of Langerhans, they supply essential elements like calcium, Zinc, Magnesium and copper for beta cells.[15] Furthermore certain plants has efficiency to prevent the conversion of starch molecules to glucose and also increase the capacity of bowl movement by proper digestion and excretion of urea.[16] Then these compounds also has capacity to scavenge the reactive oxygen and free radicles so that they also act as potent anti-oxidising agent.In this present study the biologically active compounds present in the extracts interact with carbohydrate-hydrolyzing enzymes and promote antidiabetic properties. The observed variations in chemical composition of FESAME polyherbal formulation from different plants are not only due to the type of species and also to the selected part of plant and polarity of extraction solvents.[17]

Many plant tissue and its organs like seeds, stems, roots, cotyledons, vascular tissue and pollen has ability to produce β-galactosidases enzyme. This reduces β-D-galactosyl residues from β-D-galactosides.[18] This reduction process helps in removal of carbohydrates in intestine and decreases the sugar levels. Hence this mechanism plays major role and it is the important key point to use poly herbal formulation to control diabetes. Beta galactosidase assays results showed (Fig.3) the antidiabetic efficiency of FASAME polyherbal formulation.

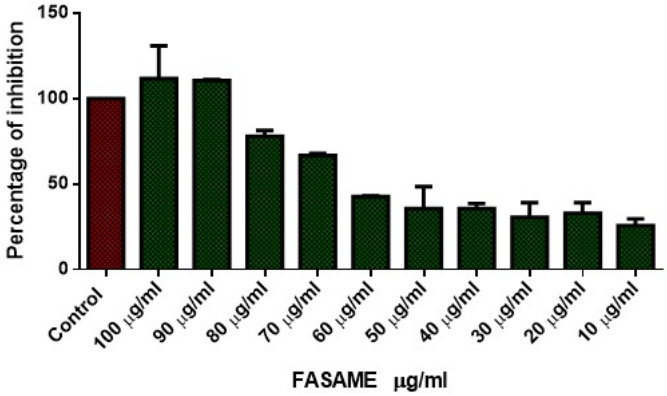

Figure 3. Percentage of inhibition of FASAME by β- galactosidase Assay

Finally, the toxicity analysis of polyherbal formulation using zebra fish larvae carried out, the results showed that the FASAME sample was not found to toxic up to the concentration of 100 µg/ml. But in 250 and 500 µg/ml there was mortality in fish larvae. Zebra fish is commonly known as Danio *rerio*, it is a small size fish with 2 to 4 cm length. It is a fresh water animal with short life cycle and so it is often used as experimental animals in research laboratories. It is largely used for toxicological studies and has 80% genetic similarity with humans.[19] Because of these reasons this fish is selected for toxicity studies. The IC_{50} was calculated and it was found to be 74.39µg/ml. The zebra fishes were exposed to different concentration. Finally, the toxicity analysis of polyherbal formulation using zebra fish larvae (Fig 4) carried out, the results showed that the FASAME sample was not found to toxic up to the concentration of 100 µg/ml. Finally, these zebra fish is proved to be good model to study toxicity of polyherbal formulations. Thus, to avoid toxicity dosage of phytochemical plays important role and lower dosage is preferred to be safe for human consumption.

DOI:10.47743/jemb-2021-55

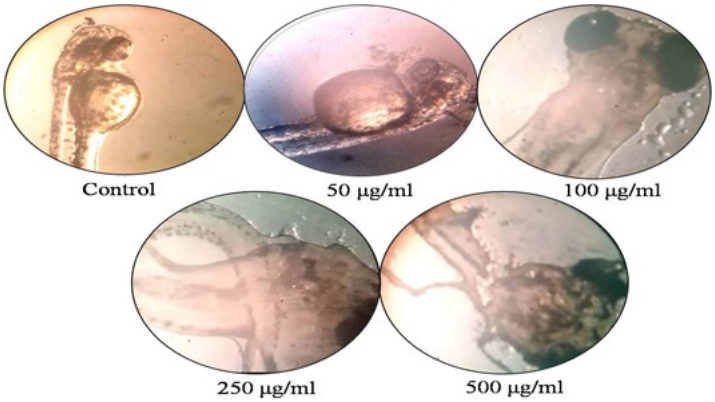

Figure 4. Effect of polyherbal formulation using zebra fish larva

Further studies should be applied to find the target compounds in these formulation using high throughput technologies.

CONCLUSIONS

In the current study the phytochemical analysis of FASAME drug showed a mixture of phytochemicals such as resins, steroids, carbohydrates, flavonoids, proteins and saponins. FASAME polyherbal formulation from various plant extracts have shown the antidiabetic activity of FASAME. The polyherbal formulation IC50 value was found to be 74.39 µg/ml. In addition, toxicity analysis with Zebra fish larvae showed a substantial increase in the percentage of viability at their lower concentration. Finally, these results showed that, FASAME polyherbal formulation could be a powerful antidiabetic drug for diabetics. The present examination has opened opportunities for further research, specifically with reference to the different dose studies and development of effective formulation for diabetes. Purification of the polyherbal extracts, formulation, and its evaluation through molecular studies will be a need for the future studies.

REFERENCES

1. Bailey CJ, Day C Traditional plant medicines as treatments for diabetes. Diabetes Care 1989; 12:553–64.
2. Singh D, Singh B, Goel RK, Traditional uses, phytochemistry and pharmacology of *Ficus religiosa*: a review. J. Ethnopharma 2011; 134(3): 565–583.
3. Malhotra V, Singh S, Tandon OP, Madhu SV, Prasad A, Sharma SB, Effect of yoga asanas on nerve conduction in type 2 diabetes. Indian. J. Physiol. Pharmacol 2002; 46:298–306.
4. Kumar CH, Kumar JNS, Ishaq BM, Rani GU, Prakash KV, Antidiabetic activity of a polyherbal preparation. Pharmacol 2010; (1):780-87.
5. Daisy P, Feril G, Kani J, Evaluation of antidiabetic activity of various extracts of *Cassia auriculata linn.* bark on streptozotocin-induced diabetic wistar rats. Int. J. Pharm. Sci 2012; 4(4):312-18.
6. Rohner A, Ried K., Sobenin, Igor AS,Heiner CB, Alain JN, A systematic review and meta-analysis on the effects of garlic preparations on blood pressure in individuals with hypertension. American J. Hypertension 2015; 28(3):414–423.
7. Simmonds M and Howes M. Plants used in the treatment of diabetes, in Traditional Medicines for Modern Time—Antidiabetic Plants. A. Soumyanath, Ed., vol. 6th, pp. 19–82. United States: CRC Press/Taylor and Francis Group; 2006.
8. Srivastava S, Lal VK, Pant KK, Polyherbal formulations based on Indian medicinal plants as antidiabetic phytotherapeutics. Phytopharmacol 2012; 2:1-15.
9. Santoriello C, Zon L I Hooked, Modeling human disease in zebrafish. J. Clin. Investig. 2012, 122; 2337–2343.

DOI:10.47743/jemb-2021-55

10. Lee J G, Cho H J, Jeong Y M, Lee J S, Genetic Approaches Using Zebrafish to Study the Microbiota–Gut–Brain Axis in Neurological Disorders. Cells. 2021, 10(3);566.
11. Debiyi O O , Sofowora F A, Phytochemical screening of medical plants, *Iloyidia*, 1978, 3; 234–246.
12. Roopashree T S , Dang R, Rani R H S, Narendra C, Antibacterial activity of antipsoriatic herbs: *Cassia tora, Momordica charantia* and *Calendula officinalis* Int. J. Appl. Res. Nat. Prod, 2008,. 1 (3); 20–28.
13. Jamous R, Zaitoun S, Husein A, Qasem I, and Ali-Shtayeh M, Screening for biological activities of medicinal plants used in traditional arabic palestinian herbal medicine, European Journal of Medicinal Plants, 2015,9(1): 1–13.
14. Bachok MF, Yusof BN, Ismail A, Hamid AA, Effectiveness of traditional Malaysian vegetables in modulating blood glucose levels. Asian Pacific J. Clini. Nutri. 2014; 23(3):369-76.
15. Boxi M, Rajesh Y, Raja Kumar V, Praveen B, Mangamma K, Extraction, phytochemical screening and in-vitro evaluation of anti-oxidant properties of *Commicarpus chinesis* (aqueous leaf extract). Int. J. Pharm. Biosci 2010; 1: 537–547
16. Zohary, Daniel, Hopf, Maria. Domestication of Plants in the Old World .3rd Ed. England:Oxford University Press; 2000.
17. Jamous R, Zaitoun S, Husein A, Qasem I, Ali M, Screening for biological activities of medicinal plants used in traditional *Arabic palestinian* herbal medicine. Europ. J. Medi. Plant 2015; 9(1): 1–13.
18. Wu A, Liu J, Isolation of the Promoter of a Cotton β- galactosidase Gene GhGal l and its Expression in Transgenic Tobacco Plants. Science in China Series C: Life Sci 2006; 54:234-245
19. Arome D, Chinedu E, The importance of toxicity testing. J. Pharma. BioSci 2014; 4:146–148.

1. Department of Zoology, Holy Cross College, Tiruchirappalli, Tamil Nadu, India.
2. Department of Biotechnology and Bioinformatics, Bishop Heber College, Tiruchirappalli, Tamil Nadu, India.

* arokiyaanitamargret@gmail.com

DOI:10.47743/jemb-2021-55

Appendix

Table S1 :Phytochemical Analysis of FASAME

S.No.	Biochemical tests	Results
1.	Detection of resins	(+)
2.	Carboxylic acid	(-)
3.	Tannins	(-)
4.	Steroids	(+)
5.	Flavonoids	(++)
6.	Carbohydrates	(+)
7.	Detection of glycosidase saponification	(-)
8.	Detection of protein Bradford method	(+)
9.	Detection of phenol ferric chloride test	(-)
10.	Biuret test	(+)
11.	Saponin test	(+)
12.	Gum test	(-)

Legend : weak (+), moderate (++), strong (+++) and absent (-).

DOI:10.47743/jemb-2021-55

Table S2 IC 50 value of FASAME

log(inhibitor) vs. normalized response -- Variable slope		
Best-fit values		
LogIC50	~ 3.264	1.872
HillSlope	~ 1.495	-7.558
IC50	~ 1839	74.39
Std. Error		
LogIC50		0.008149
HillSlope		1.003
95% Confidence Intervals		
LogIC50		1.854 to 1.889
HillSlope		-9.665 to -5.451
IC50		71.51 to 77.38
Goodness of Fit		
Degrees of Freedom	0	18
R square	1.000	0.9501
Absolute Sum of Squares	0.0	1166
Sy.x		8.047
Number of points		
Analyzed	2	20

DOI:10.47743/jemb-2021-56

TOTAL POLYPHENOLS, FLAVONOIDS CONTENTS AND ANTIOXIDANT ACTIVITY OF *ROSA* SP. GENOTIPES FROM DIFFERENT ALTITUDE OF ROMANIAN REGIONS

LACRAMIOARA OPRICA[1*], CRAITA MARIA ROSU[2]

Received: 12th of August 2021 / Revised: 5th of September 2021
Accepted: 13th of September 2021 / Published: 22nd of March 2022

Keywords: total polyphenols, flavonoids, *Rosa spp*., antioxidant activity.

Abstract. Eight wild rose hip genotypes from different altitudes varying from 3m to 902m were analyzed in order to evaluate the total polyphenols, flavonoids content and the antioxidant activity. The *Rosa spp.*rosehips collected from the Northeastern and the Southeasternof Romanianwere, as follows: *Rosa canina*, *R. caesia*, *R. corymbifera*, *R. micrantha*, *R. nitidula*, *R. rubiginosa*, *R. subcanina,* and *R. vosagiaca*. In some genotypes, the level of flavonoids and polyphenol content increased with the increasing altitude while in other it was observed a decrease. Polyphenol content reached a maximum of 144.36 mg GAE/g DW in *R. rubiginosa* whereas the lowest content of 61.72 mg GAE/g DW was recorded in *R. caesia*. The highest polyphenol content was reached at altitude of 860 m. The amount of flavonoids content ranged between 7.32 mg CE/g DW and 19.45 CE/g DW in *R. caesia* and in *R. nitidula,* respectively. The radical scavenging capacities of Rosa genotypes extracts were not positively correlated with altitude, except the *R. corymbifera* extractswhere the antioxidant activity increased with the increase of altitude.

INTRODUCTION

Rose is one of the most important crops in the floriculture industry being used as cut flowers, potted plant, and garden plants as well as in the food, perfumery, and cosmetics industries for many years (Kazaz et al., 2009).The genus Rosa contains over 100 species that are widely distributed in Europe, Asia, the Middle East and North America (Nilsson, 1997).

In Romania, there are known 29 spontaneous and subspontaneous species of *Rosa* L. genus (Oprea, 2005), widely spread throughout the country, from the sea level to the high altitude.

Rosa canina L.is a member of Rosaceae family. The fruits are well known as a rich source of vitamin C and polyphenols (Tumbas et al., 2012,Oprica et al., 2015), being considered to have one of the highest vitamin C content (30–1300 mg/100 g) among fruits and vegetables (Ziegler et al., 1986). According to Tumbas et al. (2012), vitamin C and flavonoids are responsible for the antioxidant activity of rosehips tea, while only polyphenols contribute to its antiproliferative activity. In addition, fruits of the *Rosa sp*. species, especially rose hip, contain vitamins and minerals, carotenoids, tocopherols, phenolic compounds, flavonoids, fruit acids, tannins, pectin, sugars, organic acids, amino acids and essential oils (De Vries 1980, Razungles et al. 1989, Chai and Ding, 1995, Demir and Özcan 2001; Kadakal et al. 2002, Ercisli S., 2007, Bucsa et al., 2013, Oprica et al., 2016). Because of its significant nutritional and therapeutic benefits, members of *Rosaceae* have been used both traditionally and for medicinal purposes. The bioactive compounds of the rose hips are known to have antibacterial, antidiabetic, anti-inflammatory, antiviral, immunomodulatory, antioxidant, antiulcerogenic and antidiarrheal properties as well as, effects on tumor cells andkidney stones (Orhan and Hartevioğlu, 2013). The composition and distribution of nutrients and high-valuecomponents**,** such as phenolics**,** mainly depends upon genotypes, fruit tissue, as well as, the maturity levels of fruits and to a smaller extent on environmental aspects (Manzoor et al., 2012). By the other hands, ecological factors other than altitude, soil type, temperature, and precipitation, might affect the synthesis and turnover of secondary compounds (Nobel, 1991). In this respect, some factors as seasonality, circadian rhythm, plant development, phenology, temperature, altitude, water availability, UV radiation, nutrients, pollution, mechanical stimuli and attacks by herbivores or pathogens are considered to affect the occurrence of plant metabolites (Harborne, 1993). For example, the fruits of rosehips collected in early autumn have the highest content of vitamin C compared with those collected in the late autumn. Instead, the antioxidant capacity and total phenolics of rose hips have the lower content in the early autumn due to lower levels of colored polyphenols (Pogačnik and Poklar, 2011).

With regard with the harvesting altitude of the rose hips, Rosu et al. (2011) founded that the vitamin C, total sugars and carotene content varied mostly with *Rosa sp*. genotype and only the carotene level appears to be positively correlated with altitude.

In recent years, the attention was focused on natural antioxidants of plants which have been shown to have multiple benefits on human health and nutrition.There is a lot of interest in the antioxidant activity of flavonoids and plant

DOI:10.47743/jemb-2021-56

phenolic compounds due to their potential in health promotion and disease prevention. For this reason, this study was designed to measure the contents of total polyphenolsand flavonoids as well as the antioxidant activity of eight rosehip genotypes (*Rosa canina, R. caesia, R. corymbifera, R. micrantha, R. nitidula, R. rubiginosa., R. subcanina* and *R. vosagiaca*) collected from the Northeastern and the Southeastern regions of Romania, from altitudes between 3 to 902 m.

MATERIALS AND METHODS

Sample collection and processing

Phytochemical analysis was performed on samples of rose hips representing 8 species of *Rosa spp*., as follow:*Rosa canina* L.S Str., *R. caesia* Sm., *R. corymbifera* Borkh., *R. micrantha* Sm., *R. nitidula* Besser, *R. rubiginosa* L., *R. subcanina* (Christ) Vuk., and *R. vosagiaca* N.H.F. Desp., and identified (Rosu, 2011).Samples were collectedfrom spontaneous flora of Northeast and Southeast Romanian regions from different altitudes ranging from 3 to 902 m (Table 1). Rose hips were harvested from the end of September to mid of October depending on the ripening period. The fruits were picked at the fully ripe mature stageas judged by their colour and the average samples (3 replicates) were randomly chosen from 100 fruits. The selected samples were mixed together for a homogenous distribution. Chemical composition (flavonoid and total polyphenol) and DPPH activity of dry powder obtained from whole wild rose hips fruits (seeds together with the fruit flesh) were analyzed.

Reagents. All the reagents (Folin-Ciocalteau, 2,2-diphenyl-1-picrylhydrazyl or DPPH, aluminium chloride, sodium nitrite, gallic acid and catechine) and solvents used were purchased from Sigma-Aldrich (Madrid, Spain), Fluka (Buchs, Switzerland) and Merk (Darmstadt, Germany).

Preparation of plant extract

Extraction was performed with methanol and about 0.02g of dry rose hip samples were homogenized with 80% methanol. Then it was stirred for 30 minutes and centrifuged at 3000 rpm for 15 minutes at 4°C. The supernatants were used for the further determinations.

Table1. The *Rosa spp*. fruits collected from spontaneous flora of North East and South East Romanian regions.

Taxa under study	*Sampling sites/Altitude (m)*
R. nitidula	Gradinita 1 (902m), Vatra Dornei (807m), Rusca 1 (777m), Rusca 4 (774m), Bistrita 1 (770m), Bistrita 3 (690m), Agigea 1 (3m)
R. caesia	Sadova 1 (885m)
R. vosagiaca	Sadova 6 (884m), Dorna Candreni 2 (830m)
R. subcanina	Sadova 4 (880m), Bicaz Chei 1 (685m), Agigea 2 (12m)
R. canina	Sadova 5 (875m), Dorna Candreni 1 (820m)
R. rubiginosa	Ceahlau 1 (860m)
R. micrantha	Dorna Candreni 4 (850m), Bicaz Chei 3 (668m), Sucevita 2 (595m)
R. corymbifera	Dorna Candreni 3 (840m), Sucevita 1 (691m), Dorna Candreni 5 (830m) Agigea 3 (12m)

Total polyphenols assay

The total polyphenols content was determined by using of modified Folin-Ciocalteau method (Singleton et al., 1999). The appropriately diluted sample was mixed thoroughly with Folin-Ciocalteau reagent.After four minutes, 15% Na_2CO_3 was added. The absorbance of resulting bleu-colored solution was read at 765 nm after two hours, against the blank (distilled water). The amount of the total phenolic content was expressed as milligram gallic acid equivalent per g of dried weight (mg GAE/g DW) (R^2=0.99). Three readings were taken for each sample and the results averaged.

Total flavonoids assay

The flavonoids contentwas measured following a spectrophotometric method (Dewanto et al., 2002). Briefly, methanol extracts were appropriately diluted with distilled water. Initially, 5% $NaNO_2$solution was added to each test tube; at five minutes, 10% $AlCl_3$ solution was added and then at six minutes 1.0 M NaOH was added. Finally, water was then added to the test tube and mixed well. Absorbance of resulting pink-colored solution was read at 510 nm against the blank (distilled water). The flavonoids content was expressed as mg catechin equivalents per g of dry weight (mg CE/g DW) (R^2 =0.97). Three readings were taken for each sample and the results averaged.

DPPH Free radical scavenging activity

DOI:10.47743/jemb-2021-56

The DPPH radical scavenging capacity of each extract was determined according to the method of Molyneux (2004) modified by Shirwaikar et al. (2006). DPPH radicals have an absorption that is maximal at 517 nm and which disappears with reduction by an antioxidant compound. The DPPH solution in methanol 0,1mM was prepared daily and 2ml of this solution was mixed with 20 µl of the methanol plant extracts. The control (without any antioxidant) contained 80% methanol and DPPH solution. The decrease in the absorbance of the formed blue to violet reagent (product) was determined after 20 min at 517 nm and the percentage (%) of inhibition activity was calculated using the following formula:

DPPH free radical scavenging activity (%inhibition) = (1-AE/A0)x100

where *AE* is the absorbance of the sample with extract; *A0*is the absorbance of DPPH solution with ethanol.

Statistical analysis

All experiments were carried out with three independent repetitions, the results were calculated as means ± standard errors (SE) and differences between means were assessed using ANOVA test.

RESULTS AND DISCUSSION

The amount of metabolites present in a given plant may be influenced by biological and environmental factorsas well as by biochemical, physiological, ecological and evolutionary processes (Harborne, 1993). Altitude, among other external factors, has also an effect on the contents of secondary metabolites in higher plants. In addition to incurring many climatic differences, altitude influences the quality of radiation. Especially, UV-B radiation is high in alpine sites compared with lower habitats (Barnes et al, 1987). Moreover, Zidorn (2010) showed that enhanced UV-B radiation is probably not the key factor inducing shifts in the phenolic composition in *Asteraceae* studied plants growing at higher altitudes but it is rather the temperature which decreases with altitude.

Total polyphenol content

The total polyphenol content, expressed as gallic acid equivalents (GAE) of the methanolic extracts ofeight rosehips wild genotypes, showed a great variability (Figure 1).Thus, in *R. nitidula* fruits the total polyphenol contents varied in a range of 83.02 GAE/g DW to 142.83 GAE/g DW, even the fruits were harvested from two very close altitudes (807m and 774m, respectively). Among all analyzed genotypes, the *R. rubiginosa* revealed the maximum total polyphenol content of 144.36±4.55 mg GAE/g DW in rosehips, while in *R. caesia* was measured the minimum value of 61.72±4.04 mg GAE/g DW. Both *Rosa sp.* genotypes were collected also, from close altitudes of 860m and 885m, respectively.

The polyphenol content in *R. micrantha* and *R. canina* rosehips increased with an increase in altitude from 106.04 mg GAE/g DW (at 595m) to 114.92mg GAE/g DW (at 850m) and from 95.05 mg GAE/g DW (at 820m) to 100.38 mg GAE/g DW (at 875m), respectively. Moreover, the same trend was observed in *R. corymbifera* rosehips, and the content varied from 96.28 mg GAE/g DW (at 12m) to 110.27 mg GAE/g DW (at 840m). On the other hands, in other genotypes as *R. subcanina* and *R. vosagiaca,*the polyphenol content diminished with the increase of altitude from 119.73 mg GAE/g DW (at 12m) to 77.96 mg GAE/g DW (at 880m) and from 103.67 mg GAE/g DW (at 830m) to 95.66 mg GAE/g DW (at 884m), respectively.

Long-term consumption of polyphenol-rich fruits and vegetables suggests protection of human health according to epidemiological studies and associated meta-analyzes, which is why these secondary metabolites are of high scientific interest (Pandey and Rizvi, 2009, Oprica, 2016).

DOI:10.47743/jemb-2021-56

The antioxidant capacity of phenolic compounds is mainly due to their redox properties. For this reason, they are believed to be the major phytochemicals responsible for the antioxidant activity of plants (Somaye et al., 2012). At higher altitudes, the higher solar radiation has an impact on secondary metabolite profiles. Turunen and Latola, 2005 reported an increase in phenolic compounds with increasing altitude as a response to increasing UV radiation. The authors showed that the alpine timberline plants are generally, adapted to UV-B, but on the other hand, alpine timberline plants of northern latitudes may be less protected against increasing UV-B radiation than plants from more southern latitudes and higher elevations. Meanwhile, some authors described a positive correlation (R^2= 0.55) between altitude and total polyphenols content in wild bush tea (*Athrixia phylicoides* DC.) (Nchabeleng et al., 2012) and *Hedychium spicatum* Buch. (Sandeep et al., 2011). Giorgi et al., (2010) conducted a detailed evaluation regarding the effect of environmental growth conditions on the antioxidant capacity and total phenolic content of *Achillea collina* collected from two different altitudes (600 and 1050 m). They found that growing at high altitude may constitute an effective way to significantly enhance of yarrow composition quality for both medicinal and nutritional uses.

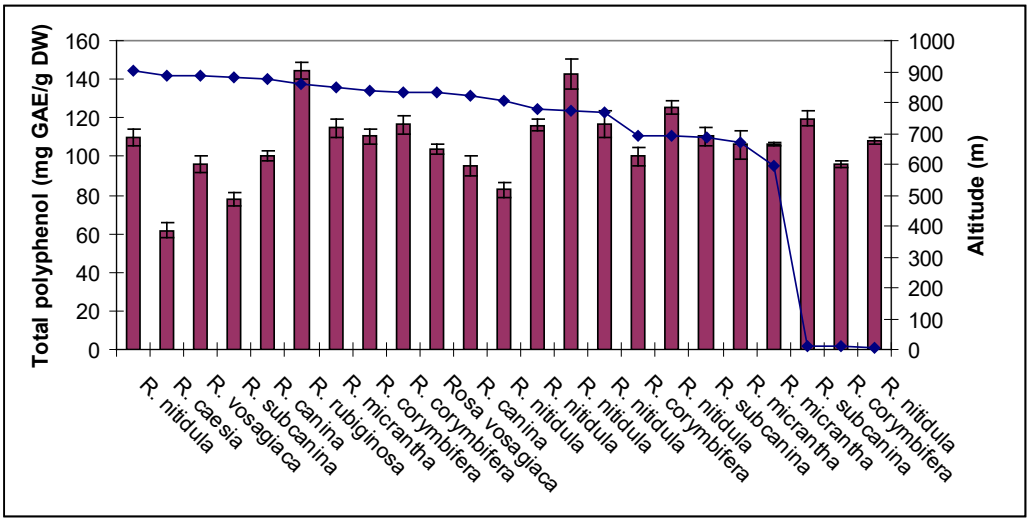

Figure1. Variation of total polyphenol content in eight *Rosa* genotypes collected from different altitude of Romania regions.

Fascella et al. (2019) reported a significant variability in bioactive compounds and antioxidant activity among the rose species collected from spontaneous Sicilian flora. The highest total polyphenol contents (6784.5 and 6241.2 mg GAE/100 g DW, respectively) and highest antioxidant activities were found in *R. canina* and *R. sempervirens* rose hips.

Many studies have reported that rose hips are rich in biologically active compounds (Ercisli, 2007; Chrubasik et al. 2008; Demir et al., 2014; Cunja et al., 2016) with positive effect on health. Barros et al. (2011), also reported the phenolic were the major antioxidant components (247.03-701.65 mg GAE/g 393 of extract). Recently, Koczka et al. (2018) identified the highest total phenolic content (766.0 mg GAE/100 g DW) and antioxidant capacity in ethanolic extracts

DOI:10.47743/jemb-2021-56

of *R. spinosissima* followed by *R. canina, R. rugosa* and, *R. gallica* and confirmed their potential as functional foods.

Flavonoids content

The content of flavonoids, expressed as catechin equivalent (CE), varied among the wild rosehipsfrom 7.32 to 19.45 mg CE/gDW (Figure 2). The highest content was measuredin*R. Nitidula* rosehips harvested from 774m, while the lowest value was noticed at *R. caesia* from higher altitude (885m).The flavonoids content varies within the genotype and also between genotypes depending on altitude. Therefore, in some genotypes, the content of flavonoids increased (*R. canina, R. micrantha* and *R. corymbifera*) with increasing altitude while in others it was observed a decrease (*R. subcanina*). Between genotypes, the highest content of flavonoids was observed in *R. corymbifera* at 830 m (14.51 mg CE/g DW) while the lowest content at 820m (10.85 mg CE/g DW). With regard to the genotypes *R. micrantha* and *R. canina*the flavonoids content increased with increasing altitude from 12.15 mg CE/g DW (595m) to 15.47 mg CE/g DW (850m) and from 11.49 mg CE/g DW (820m) to 12.53 mg CE/g DW (875m), respectively.In the case of *R. subcanina* genotype, the flavonoids content ranged between 15.54 and 9.38 mg CE/g DW at altitudesof 12m and 880m, respectively. Moreover, in some genotypes like *R. vosagiaca* picked up from the two altitudes, the content was almost identical (12 mg CE/g DW). At *R. nitidula,* between all wild rosehips collected was remarked a no uniform variation of flavonoids content depending on the altitude (902m and 3m) which ranging between 12.02 and 19.45 mg CE/g DW.

Depending on the standard used, the literature data mentioned that flavonoid content for *R. canina* was 14.71 mg/100 g extract (Daels-Rakotoarison et al. 2002), 0.33 ± 0.01 mg RE/ml (Ghazghazi et al., 2010) and 23.6 ± 4.2 mg quercetin/g extract (Montazeri et al., 2011).

In our work, the observed interspecific differences were significant in both of polyphenol (Fig. 1) and flavonoids contents (Fig. 2). Flavonoids have indeed the capacity to absorb the most energetic solar wavelengths (UV-B and UV-A), inhibit the generation of reactive oxygen species (ROS) and then quench ROS once they are formed (Brunettiet al., 2013).

In Romania, Roman et al. (2013) evaluated the amount of total phenols and total flavonoids in eight rose hip extracts from wild Transylvania populations. They reported a total polyphenols content from 575 mg/100 g frozen pulp (var. transitoriaf. ramosissima) to 326 mg/100 g frozen pulp (var. lutetianaf. fallens) and the highest value of total flavonoids of 163.3 mg/100 g frozen pulp (var.assiensis) of *Rosa canina* L. Also, they found a correlation (R^2=0,802) between the ascorbic acid content of several *Rosa canina* L. biotypes and the altitude that suggests that the content in vitamin C increase with altitude.

Also, Soare et al. (2015) determined a total phenolic content between 35.43-48.07 mg GAE/g, the antioxidant activity of maximum 363.64 mTE/100 g sp. and a flavonoid content of maximum 672.67 mg/100g in Rosehips genotypes from the spontaneous flora of Oltenia (Romania).

In a phytochemical study about *Artocarpus gomezianus*fruits collected from different altitudes of Central Western Ghats (Krishnamurthy and Sarala, 2013),the screenings of secondary metabolites revealed the presence of alkaloids, phenols, flavonoids, tannins, steroids and saponins in the fruit samples of all the regions. There was shifting among the flavonoids and phenols in the middle lower and middle higher altitudes. In contrast to phenols, tannins, and

DOI:10.47743/jemb-2021-56

steroids, the concentration of flavonoids was more in the middle and lower in higher altitudes but lowest in higher and lower altitudes.

Monschein et al. (2010) described the phytochemical compounds of *Calluna vulgaris* collected from different altitude. They found that within phenolic compounds, flavonols showed significant differences in samples collected at different altitudes with increased levels of quercetin glycosides at higher altitudes whereas no significant correlation could be found for caffeoyl quinic acids. Bernal et al. (2013) showed that the overall amount of phenolic acids and neolignan of entire leaves of *Buxus sempervirens* L. increased with altitude while the total amount of flavonoids in leaf cuticles decreased. Another study revealed that the total flavonoid content of *Hypericum perforatum* L. was positively correlated with altitude (Badgonaite et al., 2007, Tekel'ova, 2000).

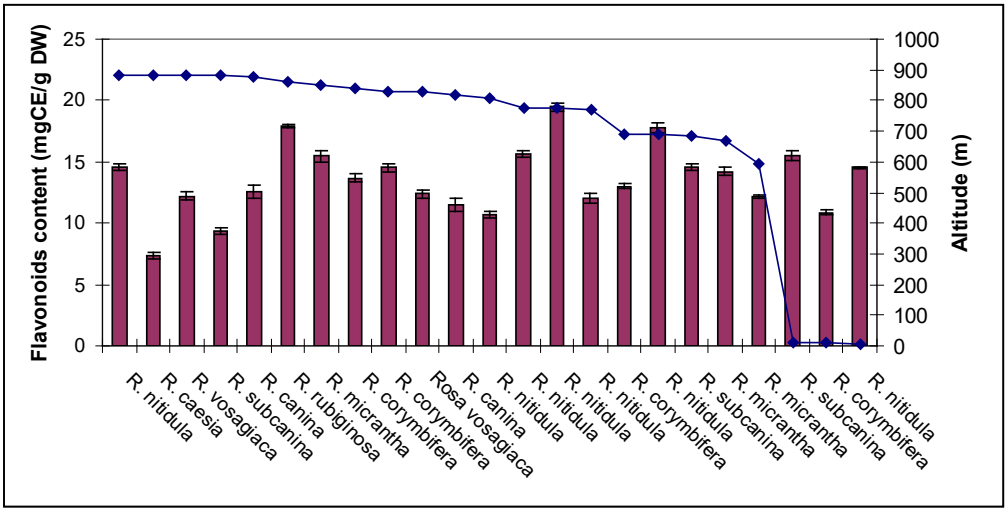

Figure 2. Variation of total flavonoids content in eight *Rosa* genotypes collected from different altitude of Romania regions

DPPH free radical scavenging activity

DPPH is a compound that possesses a nitrogen - free radical and is readily destroyed by a free radical scavenger. For this reason, it was tested the ability of the antioxidant compounds of *Rosa* genotypes who functioning as proton radical scavengers or hydrogen donors.

The screening of methanolic extractsfrom *Rosa* genotype taken into study revealed a wide variation. The DPPH free radical scavenging capacitiesexpressed as (% inhibition) is ranging from 35.44±0.22 % to 77.87±0.05 % (Table 2). The strong inhibition was also observed for the methanolic extract of *R. nitidula* while the lowest inhibition was observed at *R. subcanina*. Among all those seven *R. nitidula* collected, maximum DPPH radical scavenging percentage was 77.87±0.05% at 774m altitude while the minimum was 48.65±0.11% for genotypes collected from 807m. At this species were not noticed correlations between antioxidant activity and altitude, the values being variable.

DOI:10.47743/jemb-2021-56

At *R. corymbifera* it was observed an increase of antioxidant activity with the increasing altitude from 53.82±0.39% (at 12m) to 73.97±1.11% (at 830m). On the contrary, at *R. subcanina* the antioxidant activity increased with the decrease of altitude 61.79±0.05% (at 12 m) to 43.37±0.22% (at 880m).

Table 2. Relationship between DPPH radical-scavenging ability and altitude in eight *Rosa* genotypes collected from Romania regions

Taxa under study	*Sampling sites*	*Altitude (m)*	*DPPH free radical scavenging activity (% inhibition)*
R. nitidula	Gradinita 1	902	68,07±0,028
R. nitidula	Vatra Dornei	807	48,65±0,11
R. nitidula	Rusca 1	777	68,25±0,27
R. nitidula	Rusca 4	774	77,87±0,05
R. nitidula	Bistrita 3	690	77,79±0,16
R. nitidula	Bistrita 1	770	54,85±0,27
R. nitidula	Agigea 1	3	63,24±0,44
R. caesia	Sadova 1	885	35,44±0,94
R. vosagiaca	Sadova 6	884	54,65±0,55
R. vosagiaca	Dorna Candreni 2	830	63,64±0,11
R. subcanina	Sadova 4	880	43,37±0,2
R. subcanina	Bicaz Chei 1	685	61,31±0,055
R. subcanina	Agigea 2	12	61,79±0,05
R. canina	Sadova 5	875	65,93±0,78
R. canina	Dorna Candreni 1	820	54,06±0,27
R. rubiginosa	Ceahlau 1	860	76,81±0,44
R. corymbifera	Dorna Candreni 3	840	69,87±0,22
R. corymbifera	Dorna Candreni 5	830	73,97±1,11
R. corymbifera	Sucevita 1	691	66,83±0,055
R. corymbifera	Agigea 3	12	53,82±0,39
R. micrantha	Dorna Candreni 4	850	71,6±0,22
R. micrantha	Bicaz Chei 3	668	54,81±0,33
R. micrantha	Sucevita 2	595	61,23±0,16

Antioxidants are secondary metabolites and their contents in plants depend on varied stress conditions of vegetation (Verpoorte et al. 1999). Antioxidant activity depends on the manner in which the extracts are prepared. For this reason Buřičová and Réblová (2008), found for *R. canina* different values of antioxidant activity using DPPH radical depending on the water (62.7 mg/g) or ethanol plant extracts (6.3 mg/g). On the other hand, Wenzig et al. (2008) found that the radical scavenging activity of the methanolic extracts of *Rosa canina* was correlated very well with their total phenolic content, while ascorbic acid contributes only little to the radical-scavenging activity due to its low concentration present in the extracts.

In a study regarding the scavenging activity on DPPH radical on different concentration of *R. canina* infusion, Kilicgün and Altiner (2010) reported that the activity increased

DOI:10.47743/jemb-2021-56

significantly as a result of increasing concentration. Meanwhile, the scavenging activity dramatically decreased at higher concentrations suggesting that the same plant that optimised antioxidant capacity may also act as a prooxidant in different test systems, depending on its concentration.

The relation between the polyphenols content (mg GAE/g DW) and the antioxidant capacity (% inhibition) was determined by using linear correlations. There was a good linear correlation ($R^2 = 0.896$) between the total polyphenols content and the scavenging radical of rose hip methanolic extract (Fig. 3). So, the radical scavenging capacity of each extract could be related to their concentration of phenolic hydroxyl groups. Our results were in accord with those of Roman et al. (2013) which found a correlation of the polyphenolic compounds to DPPH in rose hip fruits and a close values of coefficient $R^2 = 0.713$.A good correlation was also found, between the flavonoids content (mg catechin/g DW) and the radical scavenging capacity ($R^2 = 0.883$) (Fig. 4).

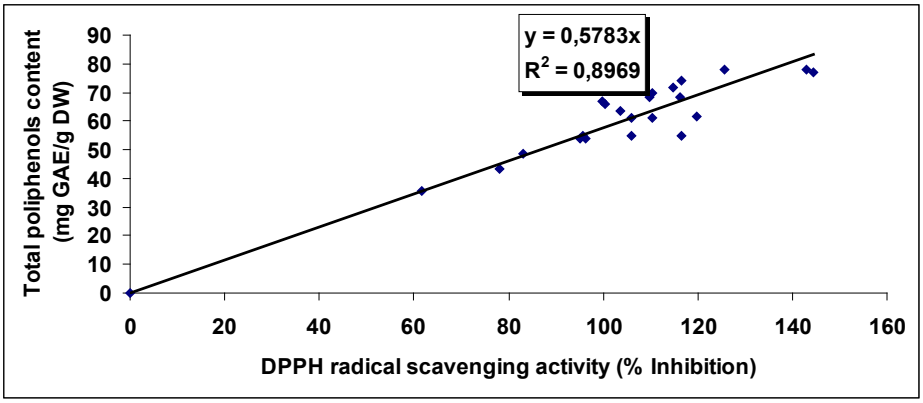

Figure 3.Correlation between total polyphenols and DPPH radical scavenging activity

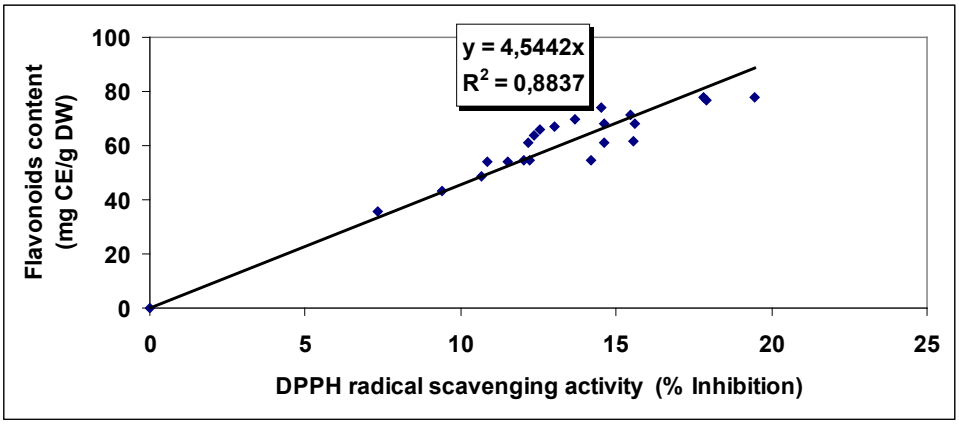

Figure 4.Correlation between flavonoids content and DPPH radical scavenging activity

DOI:10.47743/jemb-2021-56

CONCLUSIONS

The altitudinal variations of bioactive compounds are not easily studied due to the variety of possible additional factors, which can also lead to secondary metabolites profile variation.

The antioxidant compounds and antioxidant activity of eight eight rosehip genotypes collected from different altitudes of Northeast and the Southeast Romania regions have registered a large variation from 3m to 902m. The polyphenol and flavonoids content increased with the increase of altitude in the same genotype (*R. canina, R. micrantha,* and *R. corymbifera*). Regarding the decrease of polyphenol content with the increase of altitude, there were identified two genotypes *R. vosagiaca* and *R. subcanina*. In case of decrease of flavonoids content with the increase of altitude, there was only *R. subcanina*.

Generally, these wild rosehips fruits had high antioxidant capacities with an average of inhibition around 62%.The radical scavenging capacities of Rosa genotypes extracts were not positively correlated with altitude than at R. *corymbifera* where it was observed an increase of antioxidant activity with the increase of altitude.

REFERENCES

Badgonaite, E., Valdimaras, J., Ivanauskas, L., Labokas, J.,(2007): Ex situ studies on chemical and morphological variability of *Hypericum perforatum* L. in Lithuania. Biologija 53:63-70.

Barnes, P. W., Flint, S.D., Caldwell, M.M., (1987): Photosynthesis damage and protective pigments in plants from a latitudinal arctic/alpine gradient exposed to supplemental UV-B radiation in the field.Arct. Arctic Alp. Res., 19:21-27.

Barros, A., Carvalho, A.M., Ferreira, I.C.F.R.,(2011): Exotic fruits as a source of important phytochemicals: Improving the traditional use of Rosa canina fruits in Portugal. Food Res. Int., 44:2233-2236.

Bernal, M., Lorens, L., Julkunen-Tiitto, R.,Badosa, J.,Verdaguer, D.,(2013):Altitudinal and seasonal changes of phenolic compounds in*Buxus sempervirens*leaves and cuticles.Plant Physiol. Biochem., 70:471-482.

Brunetti, C., Di Ferdinando, M., Fini, A., Pollastri, S., Tattini, M., (2013):Flavonoids as antioxidants and developmental regulators: relative significance in plants and humans. Int. J.Mol. Sci., 14:3540-3555.

Bucsa, C., Atofani, D., Oprica, L., (2013):Contributions on the biochemical composition in fruits of two *Rosa* L. taxa from the spontaneous flora. Lucrări Stiintifice, S. Hort, Univ. St. Agricole si Medicină Veterinară "Ion Ionescu de la Brad" Iasi, 56(2):41-47.

Buřičová, L., Réblová, Z., (2008): Czech medicinal plants as possible sources of antioxidants. Czech J. Food Sci., 26:132-138.

Chai, J.T., Ding, Z.H., (1995): Nutrients composition of *Rosa laevigata* fruits. Sci. Technol. Food Ind., 3:26–29.

Chrubasik, C., Roufogalis, B.D., Muller – Ladner, U., Chrubasik, S.,(2013):A systematic review on the Rosa canina effect and efficacy profiles. Phytother. Res., 22(6):725-3.

Cunja, V., Mikulic-Petkovsek, M., Weber, N., Jakopic, J., Zupan, A., Veberic, R., Stampar, F., Schmitzer, V., (2016): Fresh from the ornamental garden: hips of selected rose cultivars rich in phytonutrients. J.Food Sci., 81:369-379.

Daels-Rakotoarison, D.A., Gressier, B., Trotin, F., Brunet, C., Luyckx, M., Dine, T., (2002): Effects of Rosa canina fruit extract on neutrophil respiratory burst.Phytother. Res. 16(2):157-161.

De Vries, D.P., (1980): Breeding research on rose pigments. II. Combining ability analyses of variance of four flavonoids in F1 population. Euphytica, 29:115-120.

Demir, F., Özcan, M., (2001): Chemical and technological properties of rose (*Rosa canina* L.) fruits grown wild in Turkey. J.Food Eng., 47:333-336.

Demir, N., Yildiz, O., Alpaslan, M., Hayaloglu, A.A., (2014):Evaluation of volatiles, phenolic compounds and antioxidant activities of rose hip (Rosa L.) fruits in Turkey. Lwt-Food Sci. Technol., 57(1):126-133.

Dewanto, V., Wu, X., Adom, K.K., Liu, R.H., (2002):Thermal processing enhances the nutritional value of tomatoes by increasing total antioxidant activity, J. Agric. Food Chem., 50:3010-3014.

Ercisli, S., (2007): Chemical composition of fruits in some rose (Rosa spp.) species. Food Chem., 104:1379-1384

Fascella, G., D'angiolillo, F., Mammano, M.M., Amenta, M., Romeo, F.V., Rapisarda, P., Ballistreri, G.,(2019):Bioactive compounds and antioxidant activity of four rose hip species from spontaneous Sicilian flora. Food Chem., 289:56-64.

Ghazghazi, H., Miguel, M.G., Hasnaoui, B., Sebei, H., Ksontini, M., Figueiredo, A.C., Pedro, L.G., Barroso, J.G., (2010): Phenols, essential oils and carotenoids of *Rosa canina* from Tunisia and their antioxidant activities. Afr. J. Biotechnol., 9(18):2709-2716.

DOI:10.47743/jemb-2021-56

Giorgi, A., Madeo, M., Speranza, G., Cocucci, M.,(2010): Influence of environmental factors on composition of phenolic antioxidants of*Achillea collina*Becker ex Rchb. Nat. Prod. Res.: Formerly Nat. Prod. Lett., 24(16):1546-1559.

Harborne, J. B., (1993): Introduction to ecological biochemistry, Academic Press; New York.

Kadakal, C., Nas, S., Artik, N.,(2002): Importance for human nutrition and composition of rose hip (*Rosa canina* L.) fruit and seed. Food:111-117, (in Turkish).

Kazaz, S., Baydar, H., Sabri, E., (2009): Variations in chemical compositions of *Rosa damascena* Mill. and *Rosa canina* L. fruits. Czech J. Food Sci.: 27 (3):178-184.

Kilicgün, H., Altiner, D., (2010): Correlation between antioxidant effect mechanisms and polyphenol content of *Rosa canina.*Pharmacogn. Mag.,6(23):238-241.

Koczka, N., Stefanovits-Bányai, E., Ombódi, A. (2018): Total polyphenol content and antioxidant capacity of rosehips of some rosa species. Medicines, 5(3):84.

Krishnamurthy, S. R., Sarala, P., (2013):Phytochemical studies of *Artocarpus gomezianus*Wall. ex Trecul. var.*Lakoocha*Roxb. fruits collected from various altitudes of Central Western Ghats.Indian J. Nat. Prod. Resour., 4(4):398-411.

Manzoor M., Anwar F., Saari N., Ashraf, M., (2012): Variations of Antioxidant Characteristics and Mineral Contents in Pulp and Peel of Different Apple (*Malus domestica* Borkh.) Cultivars from Pakistan, Molecules,17:390-407.

Molyneux, P. (2004): The use of the stable free radical diphenylpicrylhydrazyl (DPPH) for estimating antioxidant activity. Songklanakarin J. Sci. Technol., 26(2):211-219.

Monschein, M.,Iglesias, Neira,Kunert, O.,Bucar, F., (2010):Phytochemistry of heather (*Calluna vulgaris*(L.) Hull) and its altitudinal alteration. Phytochem Rev,9(2):205-215.

Montazeri, N., Baher, E., Mirzajani, F., Barami, Z., Yousefian, S., (2011): Phytochemical contents and biological activities of *Rosa canina* fruit from Iran'. J. Med. Plant. Res. 5(18):4584-4589.

Nchabeleng, L., Mudau, F.N., Mariga, I. K., (2012): Effects of chemical composition of wild bush tea (*Athrixia phylicoides*DC.) growing at locations differing in altitude, climate and edaphic factors. J. Med. Plants Res., 6(9): 1662-1666.

Nilsson, O., (1997): Rosa. In P. H. Davis (Ed.). Flora of Turkey and the East Aegean Islands. Edinburgh: Edinburgh University Press 4 :106-128.

Nobel, P.S., (1991): Physicochemical and Environmental Plant Physiology. Academic Press, San Diego, CA.

Oprea, A., (2005): Lista critică a plantelor vasculare din România, Edit. Univ. „Al. I. Cuza" Iaşi: 156–159

Oprica L., (2016): Metaboliti secundari din plante. Origine, structura, functii, Ed. Univ. "Al.I.Cuza"Iasi.

Oprica, L., Bucsa, C., Zamfirache, M.M., (2015): Ascorbic Acid Content of Rose Hip Fruit Depending on Altitude. Iran. J. Publ. Health, 44 (1): 138-139.

Oprica, L., Bucsa, C., Zamfirache, M.M., (2016): Evaluation of some phytochemical constituents and the antioxidant activity in six rose hips species collected from different altitude of Suceava district. Analele Stiintifice ale Universităţii "Alexandru Ioan Cuza", Secţiunea Genetică şi Biologie Moleculară,17(1):1-11.

Orhan, D.D., Hartevioğlu, A., (2013): Kuşburnu Bitkisinin Kimyasal Bileşimi ve Biyolojik Aktiviteleri The Chemical Composition and Biological Activity of Rose Hip. Spatula DD, 3(1):23-30.

Pandey K.B., Rizvi S.I., (2009): Plant polyphenols as dietary antioxidants in human health and disease, Oxidative Medicine and Cellular Longevity 2(5):270-278.

Pogačnik, L., Poklar, Ulrih N., (2011): Influence of environmental factors on antioxidants content in Rose Hips. Bulletin of the Transilvania University of Braşov, Med. Sci., 4 (53):103-110.

Razungles, A., Oszmianski, J., Sapis, J.C., (1989): Determination of carotenoids in fruits of Rosa spp. (*Rosa canina* and *Rosa rugosa*) and of chokeberry (*Aronia melanocarpa*). J. Food Sci., 54:774-775.

Roman, I., Stănilă, A., Stănilă, S., (2013): Bioactive compounds and antioxidant activity of *Rosa canina* L. biotypes from spontaneous flora of Transylvania. Chem. Cent. J., 7:73, 2-10.

Rosu, CM., Manzu, C., Olteanu, Z., Oprica, L., Oprea, A., Ciornea, E., Zamfirache, MM. (2011): Several fruit characteristics of *Rosa* spp. Genotypes from Northeastern Region of Romania.Not. Bot. Horti. Agrobo.,39(2):203-208.

Sandeep, R., Bhatt, I.D., Rawal, R.S., (2011): Total phenolic compounds and antioxidant potential of *Hedychium spicatum* Buch. Ham. ex D. Don in west Himalaya. India, 24(4-5):574-579

Shirwaikar, A, Shirwaikar, A, Rajendran, K., Punitha, IS. (2006): *In vitro* antioxidant studies on the benzyl tetra isoquinoline alkaloid berberin.Biol. Pharm. Bull., 29:1906-1910.

Singleton, V.L., Orthofer, R., Lamuela-Raventos, R.M. (1999): Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. Methods Enzymol., 299:152-178.

DOI:10.47743/jemb-2021-56

Soare, R., Babeanu, C., Bonea, D., Panita, O.,(2015): The content of total phenols, flavonoids and antioxidant activity in rosehip from the spontaneous flora from south Romania. Scientific Papers. Series A. Agronomy, 58:307-314.

Somaye, F., Rashid, J., Siavash, H.S., (2012): Antioxidant and antiradical activities of *Rosa canina* and *Rosa pimpinellifolia* fruits from West Azerbaijan. IranJ. Plant Physiol., (2(4):523-529.

Tekel'ova, D., Repcak, M., Zemkova, E., Toth, J.,(2000): Quantitative changes in dianthrones, hyperforin and flavonoids content in the flower ontogenesis of *Hypericum perforatum*. Planta Med., 66:778-780.

Tumbas, V.T., Canadanovic-Brunet, J.M., Cetojevic-Simin, D.D., Cetovic, G.S., Ethilas, S.M., Gille, L., (2012): Effect of rose hip (*Rosa canina* L.) phytochemicals on stable free radicals and human cancer cells. J. Sci. Food Agric., 92(6):1273-81.

Turunen, M., Latola, K., (2005): UV-B radiation and acclimation in timberline plants.Environ. Pollut., 137:390-403.

Verpoorte, R., van der Heijden R., ten Hoopen, H.J.G., Memelink, J., (1999): Metabolic engineering of plant secondary metabolite pathways for the production of fine chemicals. Biotechnol. Lett., 21:467-479.

Wenzig, E.M., Widowitz, U., Kunert, O., Chrubasik, S., Bucar, F., Knauder, E., (2008): Phytochemical composition and in vitro pharmacological activity of two rose hip (*Rosa canina* L.) preparations. Phytomedicine, 15:826-35.

Zidorn, C., (2010): Altitudinal variation of secondary metabolites in flowering heads of theAsteraceae: trends and causes.Phytochem. Rev., 9 (2):197-203.

Ziegler, S.J., Meier, B., Sticker, O., (1986): Fast and selective assay of L-ascorbic in rose hips by RP-HPLC coupled with electrochemical and/or spectrophotometric detection. Planta Med., 5:383-387.

[1]Department of Biology, "Alexandru Ioan Cuza" University of Iasi, Romania
[2]Department of Experimental and Applied Biology, NIRDBS - Institute of Biological Research Iasi, Romania

* lacramioara.oprica@uaic.ro

Acknowledgments: This work was supported by the project "Improving the genetic potential and complex characterization of plant biotypes group future impact on ecological and sustainable development in horticulture." (Contract PN-II- 52-142/2008).

DOI:10.47743/jemb-2021-56

DOI:10.47743/jemb-2021-57

FINISHING THE JOB - UTILITY OF LONG-READ SEQUENCING USING THE MINION FOR BACTERIAL GENOMICS

AMADA EL-SABEH[1*]

Received: 24th of August 2021 / Revised: 27th of August 2021
Accepted: 30th of August 2021 / Published: 22nd of March 2022

Keywords: bacterial genomes; third-generation-sequencing; Oxford Nanopore Technologies MinION; hybrid assembly

Abstract: Sequencing technologies have evolved dramatically since the first two bacterial genomes were published. Currently, due to second generation sequencing, millions of bacterial genomic sequences exist, although a significantly smaller amount represent completely assembled genomes. Third generation sequencing allows the analysis of single molecules, with read-lengths that cover highly complex repetitive regions previously inaccessible by short-read sequencing. However, long-read sequencing is known for producing errors which make long-read-only genome assemblies unreliable or complex if high accuracy is important for further applications. Here, Oxford Nanopore Technology's MinION, the first handheld nanopore sequencing device, is evaluated in comparison with competing sequencing platforms. The MinION's applications, potential and limitations are reviewed, focusing on its utility for bacterial genome *de novo* or hybrid assembly.

FROM FIRST TO THIRD: OVERVIEW OF SEQUENCING TECHNOLOGIES

Over a quarter of a century has passed since the first two bacterial genomes were sequenced in 1995 using what was peak technology at the time (Fleischmann et al., 1995; Fraser et al., 1995). By 2005, the number of sequenced bacteria had increased 150-fold to around 300 and already was 100-fold higher by 2015 (Land et al., 2015). Only six years later, there were almost 46 million bacterial genomic sequences stored on NCBI. This accumulation of sequencing output has risen dramatically over the past decade due to the transition from the "first generation sequencing" technology to the next generation, title which encompasses the second and the most recent third-generation sequencing technologies.

In 1977, Frederick Sanger developed the "chain-termination" technique using radiolabelled analogues of deoxyribonucleotides (Heather and Chain, 2016). Ten years later, Applied Biosystems introduced the first automatic Sanger sequencing machine using capillary based electrophoresis and replaced the radioactive reagents with a fluorometric detection system (Heather and Chain, 2016; Liu et al., 2012). Due to these improvements, Sanger's technique became more efficient and safer, establishing itself as the first generation sequencing technology and was predominantly used until 2005 (Liu et al., 2012). Despite its high accuracy and read-lengths of up to 1 kb (Heather and Chain, 2016), Sanger sequencing was still laborious and could produce only a draft genome, albeit a high-quality one. Although most bacterial genomes were complete when publicly announced, finishing the obtained genome was challenging and required a production line in itself. Moreover, obtaining a draft genome and finishing it could sum up to $50,000 (Land et al., 2015). Hence, even though Sanger sequencing produced long and highly accurate reads, the data was, at best, limited, as proven by the 15 year-long Human Genome Project (Collins et al., 2003). Furthermore, sequencing projects were handled by the few large institutions with access to the necessary funding and technological infrastructure (Land et al., 2015). Therefore, there was increasing demand for technologies that could provide higher amounts of better-quality data that was needed to uncover the answers to the increasingly complex questions that were being asked.

This eagerly expected breakthrough in sequencing started in 2005 with the launch of several Second-Generation Sequencing (SGS) platforms, beginning with 454 (now Roche 454). A year later were introduced the Genome Analyzer by Solexa (bought by the currently widely-used Illumina) and SOLiD by Agencourt (Sequencing by Oligo Ligation Detection, bought by Applied Biosystems, now part of Invitrogen) (Liu et al., 2012). These platforms and the Sanger method alike use "sequencing by synthesis" techniques, requiring the direct activity of a DNA polymerase to obtain the desired output (comprehensively reviewed by Heather and Chain, 2016). However, these novel SGS platforms were revolutionary due to their ability to sequence, in parallel, numerous PCR-amplified DNA samples (Karlsson et al., 2015). Their high throughput made it so that nowadays, it can cost under $100 to obtain a bacterial draft genome (Balloux et al., 2018) and above $1,000 for a human genome (Schwarze et al., 2020), making sequencing a no-brainer for any research team (Land et al., 2015).

Compared to the reads generated by Sanger sequencing, those produced using high throughput SGS platforms are significantly shorter, with reads produced by Illumina having ≤150 and maximally 300 bp (Wick et al., 2017). Moreover, SGS platforms yield sequences with higher error rates (~0.1–1%) (Zagordi et al., 2010) than those obtained using Sanger sequencing. Most errors are either systemic ones (Goodwin et al., 2016; Laehnemann et al., 2016), or are a consequence of the amplification step required for sample preparation (Loman et al., 2012; Ross et al., 2013). The major downfall of the short reads is in dealing with genomic repetitive sequences – if a genome contains repeats longer than the maximal read

DOI:10.47743/jemb-2021-57

length supported by the sequencing platform, it is possible that no read will cover the repeat in the genome (Loman et al., 2012). Therefore, in order to complete a genome, more data is required to ensure the necessary coverage and a larger number of contiguous sequences (contigs) has to be closed and assembled (Gurevich et al., 2013).

Prokaryote genomes are much smaller than eukaryotic ones and contain fewer and shorter repetitive sequences, of up to 10 kbp in length. However, prokaryotes may possess plasmids that can have different copy numbers than the chromosome, thus requiring a different read depth for proper assembly, and their replicons are, generally, circular (Wick and Holt, 2021). Despite the less complex genomes of prokaryotes, short-read sequencing technologies are still lacking for finishing genome assemblies. For example, to complete the *de novo* genome assembly of *Francisella* isolates which include insertions sequence elements of varying numbers and longer than Illumina reads, scientists were faced with the impossible task of assembling tens to a couple hundred discrete contigs (Karlsson et al., 2015).

Whereas with Sanger sequencing the costs were driven by sequencing itself, with SGS, the majority of the costs were skewed towards assembling the sequences into a completely finished genome. As a consequence, an increasing number of genomes were published as incomplete drafts and contained multiple contigs of various quality levels (Land et al., 2015). Despite these shortcomings, at least when discussing microorganisms, there is a consensus that most draft genomes are of "good enough quality" for a majority of common applications (Land et al., 2015, 2014). Studies that compared finished and draft versions of microbial genomes found that information was not significantly lost from generating draft genomes using Illumina sequencing (Mavromatis et al., 2012). However, it was established that the location of antimicrobial resistance genes can be of epidemiological significance; considering that repetitive insertion sequences commonly flank antimicrobial resistance genes, an incompletely assembled genome would be useless to researchers trying to asses if the resistance gene of interest is found in the chromosome or on a plasmid (Wick et al., 2017a). Furthermore, the selection of a reference genome and a performant bioinformatic pipeline is the critical factor in accurately calling single nucleotide polymorphisms (SNPs), which is crucial for preventing and tracking the transmission of microorganisms and predicting phenotypic characteristics such as antimicrobial resistance (Bush et al., 2020).

Considering these limitations of SGS, a third generation of sequencing platforms emerged in the mid-2010s, represented by the Pacific Biosciences (PacBio) RS II system and the Oxford Nanopore Technology (ONT) MinION™. These third-generation sequencing technologies (TGS) are capable of long read (of up to tens of kilobases) real-time sequencing of individual DNA and even RNA molecules. Like Sanger and SGS platforms, PacBio RS II is also based on sequencing by synthesis, but its advantages lie in the capacity of individually monitoring DNA molecules, the real-time incorporation of fluorescently labelled nucleotides and the much lower sequencing bias (compared to SGS, not Sanger) (Eid et al., 2009; Karlsson et al., 2015; Ross et al., 2013). PacBio achieved these improvements using their SMRT approach: an adapter is ligated at either end of each input DNA molecule, which is circularized and can be sequenced several times in order to achieve a higher accuracy consensus read. However, SMRT reads have high error rates, requiring deep coverage or error correction using short-reads generated with SGS, such as Illumina. For projects targeting large genomes, the yield and high cost per base sequenced using PacBio's RS II system are prohibitive factors. Supplementary, two major downfalls hindered the wide implementation of this technology in small independent labs: the instrument's initial cost and the cost of the additional necessary infrastructure (Madoui et al., 2015).

SMALL BUT MIGHTY: THE MINION™

In 2014, Oxford Nanopore Technologies released the MinION™ to over 1000 laboratories through a beta-testing program, i.e. The MinION Access Program. Although what it delivers is very similar to the PacBio system, the MinION is distinct in several important ways. Unlike previous sequencing-by-synthesis technologies, the MinION is the first commercially available single-molecule nanopore-based sequencer. The device weighs 90 g and is just 10 cm in length, making it a handheld, highly portable device that can connect to any laptop through a USB interface. Even more, the MinION is significantly more financially accessible than any other sequencing platform on the market, library construction is simplified, preparatory PCR-amplification is not necessary and the generated high-throughput data can be acquired and analysed in real-time (Feng et al., 2015; Karlsson et al., 2015; Madoui et al., 2015).

How does it work?

The MinION contains a flow cell with an array of 2048 protein nanopores split into four channels, each with four pores and sensors. Each channel can be individually controlled by an application-specific integrated circuit and allows the simultaneous processing of up to 512 nucleic acid molecules (Cherf et al., 2012; Jain et al., 2016; Magi, Semeraro, et al., 2017). When an external voltage is applied to the flow cell, particles (nucleotides) smaller than the pore size translocate through the pore (**Fig. 1., A.)**. When shifting, the negatively-charged nucleotides pass through the nanopore, the current flowing through the pore is blocked and its signal interrupted. A sensor detects these changes in ionic current as separate

DOI:10.47743/jemb-2021-57

discrete events and depicts them in what is known as a "squiggle plot" (**Fig. 1., B.**). Using graphical models, the duration, mean amplitude and variance of the discrete events is statistically analysed as a sequence of 3-6 nucleotide long k-mers (**Fig. 1., C.**) and directly correlated with the physico-chemical properties of the target molecule (Feng et al., 2015; Jain et al., 2016; Karlsson et al., 2015; Madoui et al., 2015). The translation of the current profile into nucleotide sequence information is done in real time by various base calling software such as Metrichor™ (https://metrichor.com/technology.html), Albacore, Guppy, Scrappie or Flappie (Wick et al., 2019) .

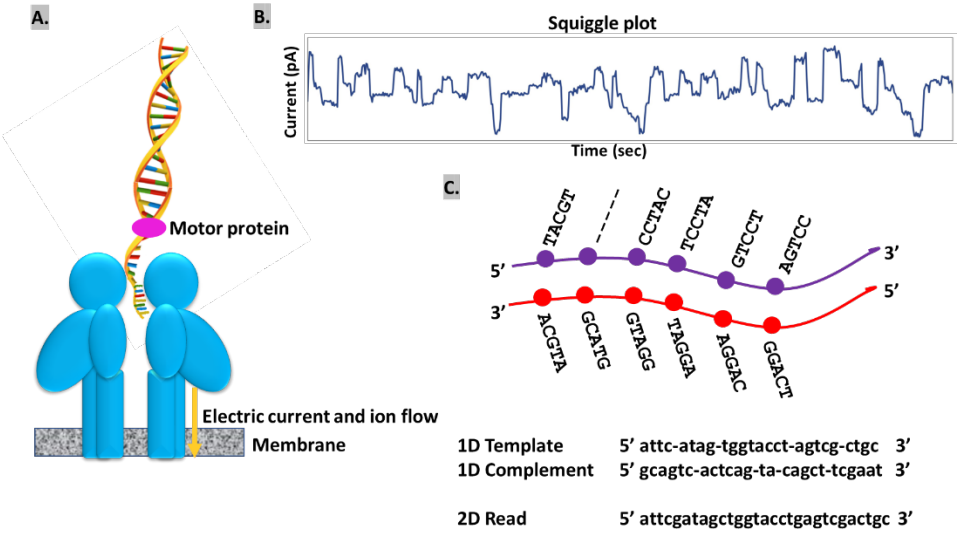

Figure 1. Flow of Oxford Nanopore Technologies MinION sequencing process

A. An adapter preloaded with a motor protein (magenta) unwinds the template strand of a duplex DNA molecule and guides it through the MinION nanopore (blue). As nucleotides pass though the pore, changes in electric current are detected by a sensor as discrete events and are depicted as a **B. "**Squiggle plot" of fluctuating electrical signals, each specific to a nucleotide's physico-chemical properties; **C.** The k-mers decoded from discrete changes in ionic current and the alignment of 1D and, optionally, 2D base calls which will provide the 2D consensus read.

For sequencing both strands of a duplex DNA molecule, adapters (preloaded with motor proteins, **Fig. 1., A.**) are ligated at either end of the DNA input (genomic or cDNA) in order to aid in strand capture, ensure the loading of a processive enzyme at the 5'-end of one strand and closely concentrate the DNA material in the nanopore's vicinity (Ip et al., 2015; Magi, Semeraro, et al., 2017). Therefore, the DNA capture rate is amplified while, on the millisecond timescale, the bound enzyme unwinds the double stranded DNA, guides a single strand through the pore and unidirectionally displaces a single nucleotide along the first (leading) DNA strand, generating the "template read" (**Fig 1., C.**). As long as the DNA strand is not damaged during sample processing, a hairpin adapter attaches at the opposite end of the DNA molecule and covalently binds together the two strands of a duplex molecule, facilitating the uninterrupted sequencing of the sense and antisense strands. After passing through the hairpin adapter, the enzyme repeats its activity on the remaining complementary strand, yielding the "complement read" which is aligned to the "template read" to generate a consensus 2D read of higher quality than 1D reads (Jain et al., 2016; Karlsson et al., 2015; Madoui et al., 2015). Using the more recent $1D^2$ chemistry which replaced the 2D method, the forward and reverse strands of a DNA duplex are sequenced without the necessity of an adapter to keep the strands in physical contact. This novel method has been proven to yield similar data quality to that obtained using the 2D chemistry, with the advantage of a simpler experimental protocol and higher number of reads produced (Tyler et al., 2018).

Several versions of the MinION chemistry and base-calling software have existed since the device's launch. The latest R9.4.1 chemistry version is based on the CsgG protein nanopore. Coupled with a computational approach that uses deep learning for base calling and the device's "fast mode", the MinION is able to sequence 450 bases/ second and yields up to 10 Gb of data from a 48-hour sequencing run (Magi, Semeraro, et al., 2017). Using circa 40 flow cells per genome, application of the R9.4 chemistry allowed the first nanopore sequencing of human genomes with a coverage of over 30× (Jain et al., 2018).

DOI:10.47743/jemb-2021-57

Benefits of using the MinION™ in lieu of other sequencing platforms

Many of the advantages of using the MinION device stem from its low cost of entry, size, portability and the ultra-long reads it produces. In addition, it can outperform other NGS platforms in several applications.

Variant detection

Called genetic variants, differences between genomes comprise of single nucleotide variants (SNVs) and structural variants (genomic alterations of at least 50 bp) such as copy number variations (mainly gene deletions and duplications), insertions, inversions, and translocations. Structural variants are responsible for more variable bases than single nucleotide variations, being involved in functional changes across populations and species and in the onset of many diseases, Mendelian and cancer alike (Magi, Semeraro, et al., 2017; Mahmoud et al., 2019). Due to the fact that an abundance of diseases is caused by mutations that can affect both genes and the non-coding genome, whole-genome sequencing has been rapidly adopted for clinical purposes because it provides a detailed set of a patient's individually specific mutation profile, allowing the customization of their treatment scheme.

Until recently, because short-read SGS platforms produce mistakes which translate into the assembled genomes, robust methods for the detection of structural variants were lacking. In 2017, using MinION's nanopore technology and the reads of up to hundreds of kilobases it can produce, Stancu et al. sequenced the whole diploid genomes of two patients at more ≥11× coverage depth. The team employed a new computational pipeline and resolved the long-range structure of a complex structural variant in the patients. Even more, the long reads obtained from MinION facilitated the identification of significantly more structural variants than those which were detected using short-read Illumina sequencing data of the same two genomes (Stancu et al., 2017).

Results obtained using the MinION for genetic discovery indicated that although the generated data was reliable enough to be useful for detecting small variants at high recall rates, sequencing errors affected the precision of the approach employed, be it genome resequencing or assembly (Magi, Semeraro, et al., 2017). This is especially significant because, contrary to expectations, errors do not occur randomly during MinION's nanopore sequencing. A characteristic of ONT's sequencing technology is that errors generally occur in homopolymers and segmentation of the current profiles are problematic in homopolymer stretches longer than 6 nucleotides, which is the maximal k-mer size detected by the pore at a time (David et al., 2017). Furthermore, errors affect bases differently, as follows: deleted bases are primarily A and T which follow A and T nucleotides, while the most frequent substitutions are of C to G and vice versa (Magi, Giusti, et al., 2017). Therefore, recurrent errors are prone to appear even at sequencing coverages of ≥30× and can lead to the discovery of false substitutions once every 10-100 kb and insertion/ deletion events every 1-10 kb. Despite this drawback and depending on the approach used for variant discovery, due to its specificity and sensitivity, MinION nanopore sequencing can be confidently used for the detection of copy number variants with high accuracy, with better results than those obtained using PacBio long-reads or SGS short-reads (Magi, Giusti, et al., 2017; Magi, Semeraro, et al., 2017).

Analysis of RNA expression

For RNA analysis on SGS platforms, several preparatory techniques are required such as fragmentation, conversion of RNA into complementary DNA and PCR amplification. These steps introduce experimental biases and impede the accurate determination of gene expression (Garalde et al., 2018; Ozsolak and Milos, 2011). These are critical issues for clinical applications such as determining resistance to certain antibiotics. For example, resistance to aminoglycosides can arise from ribonucleotide modifications such as 7-methylguanoside. When RNA is converted into cDNA, epigenetic modifications can no longer be observed. However, TGS platforms such as ONT's MinION can directly measure RNA in its native molecules without any preparatory shearing and qPCR steps and, even more, there is no limit to the length of the sequenced molecules (Smith et al., 2019). Thus far, ONT's MinION has been used to evaluate the resistome of four extensively drug-resistant *Klebsiella pneumoniae* clinical isolates (Pitt et al., 2020), estimate 3' poly(A) tail length (Krause et al., 2019; Workman et al., 2019), produce an accurate profile of the transcriptome (Bolisetty et al., 2015; Sessegolo et al., 2019), and identify novel isoforms (Byrne et al., 2017; Križanović et al., 2018) and RNA modifications (extensively referenced in Cozzuto et al., 2020).

Base modification detection

Previous to the advent of single-molecule sequencing technologies, the state-of-the-art method for identifying genome-wide DNA methylations was based on treatment with sodium bisulfite followed by SGS, which did not offer data regarding long-range methylation patterns. Nowadays, using nanopore technology, the methylation of nucleic acids can be directly detected in native molecules at the nucleotide level, in both DNA and RNA. Prior to the launch of the MinION, two groups independently showed that a single-channel nanopore system can distinguish between all five types of C-5 cytosine variants in synthetic DNA (Schreiber et al., 2013; Wescoe et al., 2014), with accuracy ranging from 92-98% for a target cytosine in a known sequence (Wescoe et al., 2014). Since then, two groups have used the MinION and developed software algorithms to identify cytosine methylation in human and/ or bacterial genomic DNA with more than 80% accuracy (Rand et al., 2017;

DOI:10.47743/jemb-2021-57

Simpson et al., 2017). Simpson et al.'s method discriminated among cytosine and 5-methylcytosine, while Rand et al.'s tool also identified 5-hydroxymethylcytosine, i.e. only two or three, respectively, of the five types of C-5 variants known. Additionally, in the case of the method employed by Simpson et al., their training set only focused on fully methylated genomic regions and did not detect those containing heterogenous methylation (Simpson et al., 2017).

However, four years after its launch, the MinION was already sensitive enough to sequence 5 picograms of purified 16S *E. coli* rRNA detected in 4.5 µg of total human RNA and to identify 7-methylguanosine and pseudouridine modifications (Smith et al., 2019). In 2020, Cozzuto et al. made available their open-source workflow for the analysis of direct RNA sequencing data, named MasterOfPores. The pipeline converts raw current intensities into processed data, maps the reads, predicts RNA modifications and estimates poly(A) tail lengths. The MasterOfPores workflow can be easily run on any computer with the Unix OS, does not require the installation of additional software and allows for four direct RNA MinION sequence data sets to be fully processed and analysed in 10 h on 100 CPUs (Cozzuto et al., 2020). Therefore, despite the current limitations, continuous efforts are being made in improving base-call accuracy and developing bioinformatic tools suitable for the accurate and thorough identification of the different base modifications in the genome and transcriptome.

Real-time targeted sequencing

For clinical applications especially, it is crucial to obtain and analyse genomic and transcriptomic data as quickly as possible. Neither Sanger nor SGS can match ONT's MinION in terms of on-site sequencing due to its portability and accessibility. The MinION can be used immediately upon arrival in an outbreak area without the need for calibration procedures or an actual laboratory set-up, which could be problematic in certain regions because of logistical (sample transportation and storage) or political issues (Lu et al., 2016). Even more, the MinION has a "Read Until" function: a mix of DNA fragments can be applied to the flow cell and when a DNA strand translocates through the nanopore, the current intensity profiles are compared to the expected pattern for a target sequence. If there is no match, that DNA strand is rejected by the nanopore, which will continue its analysis of a different DNA strand; otherwise, sequencing continues. Thus, for clinical applications such as in-filed and point-of-care, the "Read Until" function dramatically reduces the time elapsed from sample acquisition to result analysis (Jain et al., 2016). Targeted reverse transcription PCR coupled with MinION sequencing have already been used to obtain rapid data turnaround for the management of disease outbreaks such as those of the Severe Acute Respiratory Syndrome Coronavirus 2, Ebola and Zika viruses (Paden et al., 2020; Quick et al., 2017, 2016).

***De novo* assembly**

A major feature of the MinION is the read lengths it yields, which dramatically surpass those produced by the best performing SGS platforms. In its launch year, the MinION had already been used by Ip et al. for sequencing *E. coli* genomic DNA; the team obtained 1D and 2D read lengths of over 300 kb and up to 60 kb, respectively (Ip et al., 2015). In the same year, Loman and collaborators used only MinION sequence reads to *de novo* assemble the *E. coli* K-12 MG 655 chromosome into a single contig of 4.6 Mb, which had correctly ordered genes and 99.5% nucleotide identity. Instead of relying on the assemblers available at the time, which were unable to handle MinION sequencing errors, they corrected the long reads in two phases, by using a multiple-alignment process followed by polishing via a probabilistic model of the signal-level data (Loman et al., 2015). Although their approach yielded satisfactory results, the MinION produces data with high error rates compared with Sanger or SGS platforms, which have made its use problematic for *de novo* assembly algorithms designed for short reads and fewer errors (Goodwin et al., 2015). To overcome this impediment, continuous efforts have been made since the launch of the MinION Access Programme to develop computational approaches which are capable of accurately processing the high-throughput and error-prone MinION long-read sequences (Magi, Semeraro, et al., 2017). To drive this point forward, a comprehensive review of the performance of long-read assemblers for prokaryote whole genome sequencing was first published in 2019 and has been updated yearly since then (Wick and Holt, 2021).

Thus far, although ONT's nanopore technology has been primarily used for microbial sequencing, efforts have already been made to sequence and assemble several significantly more demanding eukaryotic genomes. For example, Istace et al. used the MinION to perform *de novo* sequencing and assembly of 21 genetically diverse *Saccharomyces cerevisiae* isolates and obtained assembly contiguities 14 times higher compared to Illumina-only assemblies. 65% of the chromosomes were covered by only one or two contigs, which enabled the accurate discovery and inspection of long structural variants present across the 21 sequenced genomes, variations which were generally missed using only short-read sequencing (Istace et al., 2017). Schimdt et al. used the MinION to sequence and assemble the genome of an even more complex eukaryote, namely that of *Solanum pennellii*, a wild tomato species. They obtained an assembly with an N50 value of 2.5 MB, but although the genome assembled using raw nanopore sequences was structurally highly similar to their reference genome, it was rich in homopolymer deletions and had a high error rate. Finally, the team used a hybrid assembly approach (see below) by applying Illumina short-reads to finesse the nanopore-reads assembly to an error rate lower than 0.02 when compared against the Illumina data set. Therefore, although nanopore technology and its dedicated assemblers are being continuously

DOI:10.47743/jemb-2021-57

improved, *de novo* genome assembly requires careful data analysis and polishing to obtain accurate and relevant error rates, followed by checking the genome quality and gene content (Schmidt et al., 2017).

Hybrid assembly

Because the majority of publicly-available genomes are incomplete, especially in the case of thousands of sequenced bacteria, significant interest and effort have been applied towards combining the complementary advantages of the accurate but short reads produced by Illumina sequencing and long but error prone reads resulted from TGS platforms. This method is known as "hybrid assembly" and can be achieved using one of two distinct approaches, namely short-read-first or long-read-first methods. The former method employs a scaffolding tool which uses long reads to join together Illumina contigs. The disadvantages of this approach are the misassembled sequences which arise from quite common scaffolding mistakes (Hunt et al., 2014). In long-read-first approaches either the uncorrected long reads are assembled first and short reads are used to correct errors in the assembly (Koren et al., 2017), or short reads are firstly used to error-correct long reads, the final assembly using the already-corrected long reads (Koren et al., 2012; Salmela and Rivals, 2014). Regardless if the error-correction is performed before or after the assembly, a higher long-read depth is required when the long-read-first approach is employed (Wick et al., 2017b). However, due to the read lengths produced by PacBio and ONT exceeding those of the length of repeats found in most bacterial genomes, complete hybrid genome assemblies have been achieved with even only one contig per replicon (Conlan et al., 2014; Koren et al., 2017).

THE MINION^{TM}: FOR THE LAST BRICK IN THE WALL

As previously discussed, genome assembly using only short-read sequences generally results in a number of unordered contigs. For example, an attempt by Karlsson et al. to *de novo* assemble the *Francisella* FSC996 chromosome (32% GC content) using Illumina short-reads with a 1000× coverage resulted in 40 contigs. The team changed tactics and performed hybrid genome assemblies on different *Francisella* strains using either MinION or PacBio long-reads coupled with Illumina-generated short-reads. Only a fifth of the long-reads from a single MinION run were sufficient to obtain a correct genome scaffold and easily assemble the contigs into a complete genome. Despite the older R7.3 chemistry available at that time and the higher error rates associated with it compared to the current R9.4.1, a single MinION run yielded a genome of ~99.8% sequence accuracy (Karlsson et al., 2015).

In 2017, Wick et. al used barcoded ONT libraries sequenced in multiplex on a single MinION flowcell and a hybrid-assembly approach to resolve the large genomes of 12 *Klebsiella pneumoniae* isolates (Wick et al., 2017a). The data was assembled either using ONT-only reads or a hybrid approach using complementary Illumina data, which had been previously found by the team to be insufficient for identifying the location of antimicrobial resistance genes (Gorrie et al., 2017). The group observed that ONT-only assemblies were prone to high error rates and were not substantially improved by read depth, although this parameter did positively influence sequence accuracy. The most accurate ONT-only assembly they obtained had an error rate equivalent to one error per 287 bp, implying that more often than not, a 1 kbp gene could contain an error. They concluded that such an assembly would not facilitate resistance allele or multi-locus sequence typing or be appropriate for studying phylogenomics or drug resistance transmission. Even more, they highlighted that the obtained MinION data did not accurately represent the small plasmids present in the bacterial isolates. Consequently, Unicycler, the assembly pipeline the group used, was unable to automatically complete the genome because of the missing plasmid sequence data. This issue is a great example of the importance of sample preparation depending on the scientific interest. The team hypothesized that either the DNA extraction method they used was improper for small DNA fragments such as plasmids, or the compromised plasmid recovery was a consequence of omitting the DNA shearing step during library preparation. Hence, because any existing small plasmids remained circular, no DNA strand ends were free to bind to the ONT adapters and the small plasmids evaded sequencing. Despite these caveats, the use of MinION-Illumina hybrid read sets resulted in 12 finished genomes, with a cost of around 150 USD per strain (Wick et al., 2017a).

Lemon et al. focused on testing the performance of the MinION in sequencing already isolated plasmid DNA. They resequenced three plasmids from a reference *K. pneumoniae* isolate; the accuracy of the draft genome was 99% when assembled using only MinION reads, the value increasing to 99.9% when the draft assembly was polished using Illumina MiSeq short-reads. The group also sequenced plasmid DNA from previously uncharacterized antibiotic resistant *E. coli* and *K. pneumonia* clinical isolates. The MinION reads enabled the facile detection of drug resistance genes in the draft genome assembly. Interestingly, by using isolated plasmid DNA instead of whole genomic DNA, full annotation of antimicrobial resistance genes was possible with quite low read depth, using only 2000-5000 reads, which can be produced within 20 minutes of sequencing (Lemon et al., 2017).

Even tough Wick and colleagues were able to completely assemble the large *K. pneumoniae* genomes with MinION read lengths of N50>20 kb and over 14× coverage, their previous experience indicated that bacterial strains whose genomes have more frequent and larger repetitive sequences may require more distinct approaches (Wick et al., 2017b). For example,

DOI:10.47743/jemb-2021-57

some *Shigella* (cause of bacillary dysentery) genomes contain a couple hundred pseudogenes, many high-copy-number repeats of ~1 kbp associated with hundreds of copies of insertion sequence elements and numerous indels, translocations and inversions. Their findings demonstrated that compared to *K. pneumoniae*, *Shigella* genomes require around twice the nanopore sequencing depth to obtain complete hybrid assemblies (Wick et al., 2017b; Yang et al., 2005). Another problematic species is *Acinetobacter*, with one of the earliest multiple-antibiotic-resistant isolates *A. baumannii* strain A1 containing a highly repetitive biofilm-associated gene variable in length but which can reach over 25 kbp (Holt et al., 2016). Holt et al. did assemble the genome using PacBio and Illumina reads, but this required significant manual intervention in both the assembly and annotation steps. In contrast, Wick et al. used simulated reads from the reference genome produced by Holt et al. and the hybrid assembly tool Unicycler and obtained an exclusively automatic complete genome assembly of *A. baumannii* strain A1 (Wick et al., 2017b).

Similar to Wick et al., Todd and colleagues also used sample multiplexing for short-read Illumina and long-read MinION platforms and assembled the obtained data using Unicycler. They succeeded in assembling seven genomes of *Fusobacterium* into highly accurate and singular complete chromosomes, compared to the previously available draft assemblies containing 24-67 fragmented contigs. Even more, they revealed the presence of a genomic inversion of over 450 kb in the previously existing *F. nucleatum* subsp. *nucleatum* ATCC 25586 genome assembly, which they managed to correct using the hybrid assembly approach (Todd et al., 2018).

In 2018, two independent groups at the National Microbiology Laboratory (Public Health Agency of Canada) sequenced four well-characterized isolates in replicate, using the latest flow cells, sequencing chemistries and software available at the time. When designing an experiment, consistent workflows between replicate runs are ideal. However, in this study, because of the accelerated evolution of MinION sequencing (reaction chemistries, software), the researchers were unable to ensure technical consistency throughout their experiments. Furthermore, the inter-run variability regarding the obtained yields made it challenging to estimate yield per flow cell based solely on the number of samples and input DNA. Despite these caveats, the group observed that both sequencing yield and quality had improved throughout the experiment, sequence alignment accuracies being over 94% for 1D and 2D chemistries alike, the resulting data being equally suitable for genome assembly. Overall, the high error rate (which has since been improved and is easier to overcome by the available software), the inconsistencies observed between runs caused by the rapidly changing kits, reagents and software were considered limiting factors for adopting MinION sequencing for wide-scale use. Despite this, the study recognized the advantages the MinION brings for whole genome sequencing of bacteria and its capacity for pathogen identification even in samples with DNA concentrations lower than those recommended by ONT (Tyler et al., 2018).

In a recent publication, two reference strains and two field isolates of *Campylobacter jejuni* were sequenced using Illumina MiSeq and MinION. The sequences were assembled using either designated assemblers for short-reads (SPAdes) and long-reads (Canu), respectively, or Unicycler, which performs hybrid genome assemblies using both read types. The Illumina raw data assembled using SPAdes had the most nucleotide identity and genes correctly annotated when compared to the PacBio-generated reference genomes, but the short-read lengths yielded fragmented contigs and a greater, misrepresenting number of coding sequences. On the other hand, MinION-only assemblies using Canu were contiguous and enabled the easy identification of plasmids, but had the least accuracy and contained numerous errors such as substitutions and indels, which lead to inaccurate gene annotations and sequence typing. Finally, the MiSeq and MinION data were combined to obtain hybrid genome assemblies. The number of mismatches was slightly higher than in the Illumina-only assembly, possibly because the assembly pipeline (Unicycler) heavily relies on MinION data in regions of low Illumina read coverage. The assembly accuracies were improved when the amount of MinION data used in the assembly was increased from 40× to 200× and the resulted hybrid-assembled genomes were contiguous and completely circularized. The bacterial genomes constructed using the hybrid approach were the most useful for identifying plasmids, large genomic rearrangements and repetitive elements such as genes coding for ribosomal and transport RNA (Neal-McKinney et al., 2021).

Similar findings were reported by Goldstein et al., who sequenced nine bacterial genomes with GC contents varying from low to high, using either Illumina MiSeq or the MinION. They tested short-read-first and long-read-first assembly approaches. Regarding bacterial strains with extreme GC contents, the researchers mentioned that because of the bias from the Illumina libraries, polishing the long-read data using MiSeq sequences was a challenge. In spite of this, they concluded that using MinION reads for initial assembly followed by Illumina short-reads for error-correction provided the most contiguous genomes, which were accurate enough for annotating challenging regions to sequence, such as secondary metabolite biosynthetic gene clusters and insertion sequence elements (Goldstein et al., 2019).

CONCLUSION: THE ONLY WAY IS FORWARD

The MinION is a paradigm-shifting device due to its nanopore sequencing method, portability and commercial diffusion to research and clinical applications. The use of nanopore sequencing greatly improves *de novo* genome assemblies when

DOI:10.47743/jemb-2021-57

considering N50, contig numbers and its robustness with extreme GC content organisms, allowing the detection and exploration of structural and sequence variants. Even though assembling MinION nanopore reads alone is feasible, issues with sequence accuracy and small plasmid recovery were reported. Therefore, scientists have called for more suitable approaches to library preparation and base-calling algorithms to tackle the caveats of *de novo* assemblies relying exclusively on nanopore sequences (Wick et al., 2017a). Because long-read sequencing is becoming more common especially in microbial genomics, long-read assembly is also on a continuous rise. Therefore, the development and refinement of designated assemblers is paramount for the scientific community to be able to actually use the full potential of these rapidly and dynamically-evolving sequencing technologies (Wick and Holt, 2021).

Despite the drawbacks that nanopore sequencing has for *de novo* genome assembly, it has become a robust and highly-appreciated method for succeeding where other platforms have not. Although at first glance, the MinION seems to be highly similar to PacBio in that they produce similarly sized reads, the two systems have been proven to have quite distinct applications and target users. Whereas the MinION stands out due to its convenient size, PacBio platforms are dramatically large and heavy and require a substantial initial investment, being more suitable for sequencing centres where space and infrastructure is not an issue. The MinION is the complete opposite (Karlsson et al., 2015). ONT's device has thus far been shown to be easily operated in the field for rapid real-time pathogen identification, which is advantageous for more general sample analysis than what can be achieved using real-time PCR. The device performed very well even with lower DNA input than what the manufacturer deemed necessary, as indicated by Tyler et al. Even more, the MinION has already gone where no other sequencing device has. As part of a 6-month long NASA experiment on the International Space Station, it was used to successfully sequence bacteriophage, bacterial and eukaryote DNA in microgravity. In parallel, on Earth, *de novo* assembly of the MinION data showed over 96% consensus pairwise identity. The MinION's performance was benchmarked against Illumina's MiSeq and PacBio's RS II platforms, the results being promising for MinION applications for in-space monitorization of human health and response to spaceflight and the identification of DNA-based extra-terrestrial lifeforms (Castro-Wallace et al., 2017).

Already a jack of many trades, the MinION has proven itself a robust and reliable master for microbial sequencing and assembly, especially when combined with complementary sequencing methods such as those provided by Illumina. As was discussed above, hybrid genome assembly using MinION long reads enables closing the gap regions where the reach of short-read sequencers was insufficient. Complete and accurate genomes are essential for a variety of scientific and clinical applications among which phylogenetic studies and infectious disease epidemiology, for which a hybrid assembly approach is, momentarily, the golden standard for obtaining high-quality, detailed and accurate data sets. However, a hybrid assembly approach is not always necessary and the additional labour and costs may be misplaced for bulk sequencing of bacterial isolates. On the other hand, demonstrating genetic relatedness or identity is commonly a necessity when managing pathogen outbreaks and establishing regulatory action. The most accurate comparison between isolates would be ensured by using hybrid-assembled genomes due to their completeness and contiguity, which enables the distinction of potential drug resistance genes and virulence factors belonging either to plasmid or to chromosomal DNA. If interested in the characterization of the complete genetic content of a bacterium or if wanting to compare highly related isolates, then hybrid assembly is the best method for obtaining the additional detail which may lack from incomplete short-read-only or error-prone long-read-only assemblies (Neal-McKinney et al., 2021).

To conclude, whereas other sequencing platforms require high capital investment which restricts the location of sequencing infrastructure to high-brow sequencing centres (Stancu et al., 2017), the MinION is currently unmatched in terms of range of applications, up-front costs, ease of use, initial set-up, space and infrastructure requirements. Furthermore, the MinION is an extremely flexible device and opens the genomics, transcriptomics and epigenomics field to virtually anyone with sufficient funds and knowledge; as long as a Unix-compatible computer is in sight, the MinION can be used. Although similar technologies in terms of long-read sequencing do exist, the MinION stands out because it can reach smaller, independent laboratories which may be financially unable or just undesiring of a high investment such as that required for a PacBio system. Currently, the biggest caveats of MinION sequencing continue to be the error rate, which has been declining since its launch, and the dynamic changes to the flow cells, reagents and kits. The higher error rates can be handled via a few different approaches such as bioinformatic work-arounds, increasing sequencing depth, error-correcting using already available short-read datasets or sequencing in parallel using a complementary SGS platform. Regarding the common changes to the technology, these seem to be problematic mainly for wide-scale use. Considering the timeline of sequencing technologies and although fast evolving, nanopore sequencing has just surpassed its infancy and, by all accounts, its future is bright and shining and will open many doors for us (or rather, in the case of genomes, close them).

DOI:10.47743/jemb-2021-57

REFERENCES

Balloux, F., Brønstad Brynildsrud, O., van Dorp, L., Shaw, L. P., Chen, H., Harris, K. A., … Eldholm, V. (2018, December 1). From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends in Microbiology*, Vol. 26, pp. 1035–1048. Elsevier. https://doi.org/10.1016/j.tim.2018.08.004

Bolisetty, M. T., Rajadinakaran, G., and Graveley, B. R. (2015). Determining exon connectivity in complex mRNAs by nanopore sequencing. *Genome Biology*, *16*(1), 1–12. https://doi.org/10.1186/s13059-015-0777-z

Bush, S. J., Foster, D., Eyre, D. W., Clark, E. L., De Maio, N., Shaw, L. P., … Walker, A. S. (2020). Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism–calling pipelines. *GigaScience*, *9*(2), 1–21. https://doi.org/10.1093/GIGASCIENCE/GIAA007

Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., … Vollmers, C. (2017). Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications*, *8*(1), 1–11. https://doi.org/10.1038/ncomms16027

Castro-Wallace, S. L., Chiu, C. Y., John, K. K., Stahl, S. E., Rubins, K. H., McIntyre, A. B. R., … Burton, A. S. (2017). Nanopore DNA sequencing and genome assembly on the International Space Station. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-18364-0

Cherf, G. M., Lieberman, K. R., Rashid, H., Lam, C. E., Karplus, K., and Akeson, M. (2012). Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nature Biotechnology*, *30*(4), 344–348. https://doi.org/10.1038/nbt.2147

Collins, F. S., Morgan, M., and Patrinos, A. (2003). The Human Genome Project: Lessons from large-scale biology. *Science*, Vol. 300, pp. 286–290. Science. https://doi.org/10.1126/science.1084564

Conlan, S., Thomas, P. J., Deming, C., Park, M., Lau, A. F., Dekker, J. P., … Segre, J. A. (2014). Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing *Enterobacteriaceae*. *Science Translational Medicine*, *6*(254). https://doi.org/10.1126/scitranslmed.3009845

Cozzuto, L., Liu, H., Pryszcz, L. P., Pulido, T. H., Delgado-Tejedor, A., Ponomarenko, J., and Novoa, E. M. (2020). MasterOfPores: A workflow for the analysis of oxford nanopore direct RNA sequencing datasets. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00211

David, M., Dursi, L. J., Yao, D., Boutros, P. C., and Simpson, J. T. (2017). Nanocall: An open source basecaller for Oxford Nanopore sequencing data. *Bioinformatics*, *33*(1), 49–55. https://doi.org/10.1093/bioinformatics/btw569

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., … Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, *323*(5910), 133–138. https://doi.org/10.1126/science.1162986

Feng, Y., Zhang, Y., Ying, C., Wang, D., and Du, C. (2015, February 1). Nanopore-based fourth-generation DNA sequencing technology. *Genomics, Proteomics and Bioinformatics*, Vol. 13, pp. 4–16. Elsevier. https://doi.org/10.1016/j.gpb.2015.01.009

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., … Venter, J. C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, *269*(5223), 496–512. https://doi.org/10.1126/science.7542800

Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., … Venter, J. C. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science*, *270*(5235), 397–403. https://doi.org/10.1126/science.270.5235.397

Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., … Turner, D. J. (2018). Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods 2018 15:3*, *15*(3), 201–206. https://doi.org/10.1038/nmeth.4577

Goldstein, S., Beka, L., Graf, J., and Klassen, J. L. (2019). Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genomics*, *20*(1), 1–17. https://doi.org/10.1186/s12864-018-5381-7

Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C., and McCombie, W. R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Research*, *25*(11), 1750–1756. https://doi.org/10.1101/gr.191395.115

Goodwin, S., McPherson, J. D., and McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, Vol. 17, pp. 333–351. Nature Publishing Group. https://doi.org/10.1038/nrg.2016.49

Gorrie, C. L., Mirc Eta, M., Wick, R. R., Edwards, D. J., Thomson, N. R., Strugnell, R. A., … Holt, K. E. (2017). Gastrointestinal carriage is a major reservoir of *Klebsiella pneumoniae* infection in intensive care patients. *Clinical Infectious Diseases*, *65*(2), 208–215. https://doi.org/10.1093/cid/cix270

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., A, G., V, S., … G, T. (2013). QUAST: quality assessment tool for genome assemblies. *29*(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Heather, J. M., and Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, Vol. 107, pp. 1–8. Elsevier. https://doi.org/10.1016/j.ygeno.2015.11.003

DOI:10.47743/jemb-2021-57

Holt, K. E., Hamidian, M., Kenyon, J. J., Wynn, M. T., Hawkey, J., Pickard, D., and Hall, R. M. (2016). Genome sequence of *Acinetobacter baumannii* strain A1, an early example of antibioticresistant global clone 1. *Genome Announcements*, *3*(2). https://doi.org/10.1128/genomeA.00032-15

Hunt, M., Newbold, C., Berriman, M., and Otto, T. D. (2014). A comprehensive evaluation of assembly scaffolding tools. *Genome Biology 2014 15:3*, *15*(3), 1–15. https://doi.org/10.1186/GB-2014-15-3-R42

Ip, C. L. C., Loose, M., Tyson, J. R., de Cesare, M., Brown, B. L., Jain, M., … Olsen, H. E. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Research*, *4*. https://doi.org/10.12688/f1000research.7201.1

Istace, B., Friedrich, A., D'Agata, L., Faye, S., Payen, E., Beluche, O., … Aury, J. M. (2017). De novo assembly and population genomic survey of natural yeast isolates with the Oxford Nanopore MinION sequencer. *GigaScience*, *6*(2). https://doi.org/10.1093/gigascience/giw018

Jain, M., Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., … Loose, M. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology 2018 36:4*, *36*(4), 338–345. https://doi.org/10.1038/nbt.4060

Jain, M., Olsen, H. E., Paten, B., and Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*, *17*(1), 1–11. https://doi.org/10.1186/s13059-016-1103-0

Karlsson, E., Lärkeryd, A., Sjödin, A., Forsman, M., and Stenberg, P. (2015). Scaffolding of a bacterial genome using MinION nanopore sequencing. *Scientific Reports 2015 5:1*, *5*(1), 1–8. https://doi.org/10.1038/srep11996

Koren, S., Schatz, M. C., Walenz, B. P., Martin, J., Howard, J. T., Ganapathy, G., … Phillippy, A. M. (2012). Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nature Biotechnology 2012 30:7*, *30*(7), 693–700. https://doi.org/10.1038/nbt.2280

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Krause, M., Niazi, A. M., Labun, K., Torres Cleuren, Y. N., Müller, F. S., and Valen, E. (2019). TailFindR: Alignment-free poly(A) length measurement for Oxford Nanopore RNA and DNA sequencing. *RNA*, *25*(10), 1229–1241. https://doi.org/10.1261/rna.071332.119

Križanović, K., Echchiki, A., Roux, J., and Šikić, M. (2018). Evaluation of tools for long read RNA-seq splice-aware alignment. *Bioinformatics*, *34*(5), 748–754. https://doi.org/10.1093/bioinformatics/btx668

Laehnemann, D., Borkhardt, A., and McHardy, A. C. (2016). Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. *Briefings in Bioinformatics*, *17*(1), 154–179. https://doi.org/10.1093/bib/bbv029

Land, M., Hauser, L., Jun, S. R., Nookaew, I., Leuze, M. R., Ahn, T. H., … Ussery, D. W. (2015, February 27). Insights from 20 years of bacterial genome sequencing. *Functional and Integrative Genomics*, Vol. 15, pp. 141–161. Springer. https://doi.org/10.1007/s10142-015-0433-4

Land, M., Hyatt, D., Jun, S. R., Kora, G. H., Hauser, L. J., Lukjancenko, O., and Ussery, D. W. (2014). Quality scores for 32,000 genomes. *Standards in Genomic Sciences*, *9*(1), 1–10. https://doi.org/10.1186/1944-3277-9-20

Lemon, J. K., Khil, P. P., Frank, K. M., and Dekker, J. P. (2017). Rapid nanopore sequencing of plasmids and resistance gene detection in clinical isolates. *Journal of Clinical Microbiology*, *55*(12), 3530–3543. https://doi.org/10.1128/JCM.01069-17

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., … Law, M. (2012). Comparison of next-generation sequencing systems. *Journal of Biomedicine and Biotechnology*, Vol. 2012. J Biomed Biotechnol. https://doi.org/10.1155/2012/251364

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., and Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*, *30*(5), 434–439. https://doi.org/10.1038/nbt.2198

Loman, N. J., Quick, J., and Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nature Methods*, *12*(8), 733–735. https://doi.org/10.1038/nmeth.3444

Lu, H., Giordano, F., and Ning, Z. (2016). Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics and Bioinformatics*, Vol. 14, pp. 265–279. Elsevier. https://doi.org/10.1016/j.gpb.2016.05.004

Madoui, M. A., Engelen, S., Cruaud, C., Belser, C., Bertrand, L., Alberti, A., … Aury, J. M. (2015). Genome assembly using Nanopore-guided long and error-free DNA reads. *BMC Genomics*, *16*(1), 1–11. https://doi.org/10.1186/s12864-015-1519-z

Magi, A., Giusti, B., and Tattini, L. (2017). Characterization of MinION nanopore data for resequencing analyses. *Briefings in Bioinformatics*, *18*(6), 940–953. https://doi.org/10.1093/bib/bbw077

Magi, A., Semeraro, R., Mingrino, A., Giusti, B., and D'Aurizio, R. (2017). Nanopore sequencing data analysis: State of the art, applications and challenges. *Briefings in Bioinformatics*, *19*(6), 1256–1272. https://doi.org/10.1093/bib/bbx062

Mahmoud, M., Gobet, N., Cruz-Dávalos, D. I., Mounier, N., Dessimoz, C., and Sedlazeck, F. J. (2019). Structural variant calling: The long and the short of it. *Genome Biology*, Vol. 20, pp. 1–14. BioMed Central. https://doi.org/10.1186/s13059-019-1828-7

DOI:10.47743/jemb-2021-57

Mavromatis, K., Land, M. L., Brettin, T. S., Quest, D. J., Copeland, A., Clum, A., … Kyrpides, N. C. (2012). The Fast Changing Landscape of Sequencing Technologies and Their Impact on Microbial Genome Assemblies and Annotation. *PLoS ONE*, *7*(12), e48837. https://doi.org/10.1371/journal.pone.0048837

Neal-McKinney, J. M., Liu, K. C., Lock, C. M., Wu, W. H., and Hu, J. (2021). Comparison of MiSeq, MinION, and hybrid genome sequencing for analysis of Campylobacter jejuni. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-84956-6

Ozsolak, F., and Milos, P. M. (2011). RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics*, Vol. 12, pp. 87–98. Nature Publishing Group. https://doi.org/10.1038/nrg2934

Paden, C. R., Tao, Y., Queen, K., Zhang, J., Li, Y., Uehara, A., and Tong, S. (2020). Rapid, Sensitive, Full-Genome Sequencing of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging Infectious Diseases*, *26*(10), 2401–2405. https://doi.org/10.3201/eid2610.201800

Pitt, M. E., Nguyen, S. H., Duarte, T. P. S., Teng, H., Blaskovich, M. A. T., Cooper, M. A., and Coin, L. J. M. (2020). Evaluating the genome and resistome of extensively drug-resistant *Klebsiella pneumoniae* using native DNA and RNA Nanopore sequencing. *GigaScience*, *9*(2), 1–14. https://doi.org/10.1093/gigascience/giaA002

Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., … Loman, N. J. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nature Protocols*, *12*(6), 1261–1266. https://doi.org/10.1038/nprot.2017.066

Quick, J., Loman, N. J., Duraffour, S., Simpson, J. T., Severi, E., Cowley, L., … Carroll, M. W. (2016). Real-time, portable genome sequencing for Ebola surveillance. *Nature*, *530*(7589), 228–232. https://doi.org/10.1038/nature16996

Rand, A. C., Jain, M., Eizenga, J. M., Musselman-Brown, A., Olsen, H. E., Akeson, M., and Paten, B. (2017). Mapping DNA methylation with high-throughput nanopore sequencing. *Nature Methods*, *14*(4), 411–413. https://doi.org/10.1038/nmeth.4189

Ross, M. G., Russ, C., Costello, M., Hollinger, A., Lennon, N. J., Hegarty, R., … Jaffe, D. B. (2013). Characterizing and measuring bias in sequence data. *Genome Biology*, *14*(5), 1–20. https://doi.org/10.1186/gb-2013-14-5-r51

Salmela, L., and Rivals, E. (2014). LoRDEC: Accurate and efficient long read error correction. *Bioinformatics*, *30*(24), 3506–3514. https://doi.org/10.1093/bioinformatics/btu538

Schmidt, M. H.-W., Vogel, A., Denton, A. K., Istace, B., Wormit, A., van de Geest, H., … Usadel, B. (2017). De novo assembly of a new *Solanum pennellii* accession using nanopore sequencing. *Plant Cell*, *29*(10), 2336–2348. https://doi.org/10.1105/tpc.17.00521

Schreiber, J., Wescoe, Z. L., Abu-Shumays, R., Vivian, J. T., Baatar, B., Karplus, K., and Akeson, M. (2013). Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(47), 18910–18915. https://doi.org/10.1073/pnas.1310615110

Schwarze, K., Buchanan, J., Fermont, J. M., Dreau, H., Tilley, M. W., Taylor, J. M., … Wordsworth, S. (2020). The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genetics in Medicine*, *22*(1), 85–94. https://doi.org/10.1038/s41436-019-0618-7

Sessegolo, C., Cruaud, C., Da Silva, C., Cologne, A., Dubarry, M., Derrien, T., … Aury, J. M. (2019). Transcriptome profiling of mouse samples using nanopore sequencing of cDNA and RNA molecules. *Scientific Reports*, *9*(1), 1–12. https://doi.org/10.1038/s41598-019-51470-9

Simpson, J. T., Workman, R. E., Zuzarte, P. C., David, M., Dursi, L. J., and Timp, W. (2017). Detecting DNA cytosine methylation using nanopore sequencing. *Nature Methods*, *14*(4), 407–410. https://doi.org/10.1038/nmeth.4184

Smith, A. M., Jain, M., Mulroney, L., Garalde, D. R., and Akeson, M. (2019). Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. *PLoS ONE*, *14*(5). https://doi.org/10.1371/journal.pone.0216709

Stancu, C. M., Van Roosmalen, M. J., Renkens, I., Nieboer, M. M., Middelkamp, S., De Ligt, J., … Kloosterman, W. P. (2017). Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01343-4

Todd, S. M., Settlage, R. E., Lahmers, K. K., and Slade, D. J. (2018). *Fusobacterium* genomics using MinION and Illumina sequencing enables genome completion and correction. *MSphere*, *3*(4). https://doi.org/10.1128/msphere.00269-18

Tyler, A. D., Mataseje, L., Urfano, C. J., Schmidt, L., Antonation, K. S., Mulvey, M. R., and Corbett, C. R. (2018). Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-29334-5

Wescoe, Z. L., Schreiber, J., and Akeson, M. (2014). Nanopores discriminate among five C5-cytosine variants in DNA. *Journal of the American Chemical Society*, *136*(47), 16582–16587. https://doi.org/10.1021/ja508527b

Wick, R. R., and Holt, K. E. (2021). Benchmarking of long-read assemblers for prokaryote whole genome sequencing. *F1000Research*, *8*, 2138. https://doi.org/10.12688/f1000research.21782.4

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017a). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microbial Genomics*, *3*(10), e000132. https://doi.org/10.1099/MGEN.0.000132

DOI:10.47743/jemb-2021-57

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017b). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Computational Biology*, *13*(6), e1005595. https://doi.org/10.1371/journal.pcbi.1005595

Wick, R. R., Judd, L. M., and Holt, K. E. (2019). Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biology*, *20*(1), 1–10. https://doi.org/10.1186/s13059-019-1727-y

Workman, R. E., Tang, A. D., Tang, P. S., Jain, M., Tyson, J. R., Razaghi, R., … Timp, W. (2019). Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nature Methods*, *16*(12), 1297–1305. https://doi.org/10.1038/s41592-019-0617-2

Yang, F., Yang, J., Zhang, X., Chen, L., Jiang, Y., Yan, Y., … Jin, Q. (2005). Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Research*, *33*(19), 6445–6458. https://doi.org/10.1093/nar/gki954

Zagordi, O., Klein, R., Däumer, M., and Beerenwinkel, N. (2010). Error correction of next-generation sequencing data and reliable estimation of HIV quasispecies. *Nucleic Acids Research*, *38*(21), 7400–7409. https://doi.org/10.1093/nar/gkq655

AKNOWLEDGEMENTS AND AUTHOR AFFILIATION

The author was supported by a grant from the Romanian National Authority for Scientific Research and Innovation, CNCS-UEFISCDI, project no. PN-III-P4-ID-PCE-2020-0656: Sequencing the genome of a useful bacteria: *Paenarthrobacter nicotinovorans* – next step in extending its biotechnological applications.

1 Biochemistry and Molecular Biology Laboratory, Department of Biology, Alexandru Ioan Cuza University of Iasi, Iasi 700506, Romania

* amada.elsabeh@uaic.ro