A Monte Carlo simulation for bioprospecting the endemic New Zealand terrestrial flora for antibiotic drug leads

Eric J Buenz

Nelson Marlborough Institute of Technology, Nelson, New Zealand

Correspondence to:

Eric J Buenz
Nelson Marlborough Institute of Technology
322 Hardy Street
Nelson, 7120
New Zealand

Email: Eric.Buenz@nmit.ac.nz
Tel: +64 (0)22 513 5791

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

Abstract

Background. Natural product libraries are important tools for drug discovery. However, until now, there has not been a system to allow projections of the potential number of hits from creating these libraries. The objective of this study was to develop a stochastic model system that predicts the number of hits from creating a natural product library.

Methods. A Monte Carlo simulation was developed with data from the peer-reviewed literature. Using types of endemic New Zealand terrestrial flora as examples, the number of antibacterial hits expected from creating natural product libraries were calculated.

Results. The model predicts the following bounds for the 90% range of validated antibiotic leads for the categories of the terrestrial endemic flora of New Zealand with a right skewed distribution: [grasses: 1.43-6.50; liverworts: 2.75-12.5; fungi: 45.2-207; mosses: 0.98-4.48; vascular plants: 21.4-97.8]. Furthermore, per full-time equivalent (FTE) person employed on the project, a mean of 1.31 hits (90% range 0.48-2.42) is expected.

Discussion. This model system allows the number of expected hits to be calculated when developing a natural product library for a therapeutic target. There is an opportunity to create a natural product library from New Zealand endemic terrestrial flora. This model is scalable to other geographic areas as well as to other therapeutic targets and screening systems.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

Introduction

Nearly 30% of FDA-approved drugs from 2008 to 2012 originated from natural products (Tao et al. 2014). This percentage is remarkably high, considering the deliberate commercial shift to combinatorial-chemistry products as opposed to natural products as a starting point for drug discovery (Li & Vederas 2009; Lipinski & Hopkins 2004). Extrapolation from biodiversity analyses and the endogenous characteristics of natural products indicate this natural-product resource is a future source of drugs (Zhu et al. 2012). The primary challenge continues to be that prior to evaluation of natural products for bioactivity, product libraries need to be created (Harvey et al. 2015a).

New Zealand has unique natural history (Mortimer 2004) and has a remarkably high rate of endemism. For example, 68% of the identified plants are endemic (McGlone et al. 2001) resulting in an identifiable population of 2357 endemic vascular plants (De Lange et al. 2006) potentially with more still unidentified (De Lange et al. 2009). Similarly, New Zealand has approximately >300 endemic liverworts (more than any other country), 157 endemic grasses, and 108 endemic mosses, and >5000 endemic fungi. Few of these endemic plants have undergone comprehensive evaluation for novel drug leads, although many have uses in indigenous Maori culture. This historic use can be a valuable tool (Atanasov et al. 2015; Gu et al. 2014; Gyllenhaal et al. 2012; Schwikkard & Mulholland 2014; Sucher 2013), however ethnographic information needs to be rigorously evaluated (Albuquerque et al. 2014).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

There is a need for new antibiotics (Fischbach & Walsh 2009) and natural products are a promising source of material for new antimicrobial drug leads (Clardy et al. 2006; Cragg & Newman 2013). Bioprospecting natural sources for new antimicrobials is reasonable (Gu et al. 2013), especially since the generalized characteristics of natural products are superior to combinatorial chemistry products for antimicrobials (Berdy 2012) primarily because these compounds generally adhere to Lipinski's Rule of Five (Harvey 2008). However, assessing the risk and reward of these natural-product prospecting activities can be challenging.

Stochastic models are commonly used in drug discovery and development as a technical tool in post-library-creation screening or process management (Michelson et al. 2006; Yu 2012). However, these tools have not been used to extrapolate the potential number of hits returned for generation of a new natural product library. Using stochastic models in this context allows the probability of identifying a hit from a potential natural product library to be calculated; particularly in geographies where a deliberate approach to developing natural product library has yet to occur.

Here we present a stochastic model that projects: 1) the number of potential new antibiotic hits that would be identified in a screen using endemic New Zealand terrestrial flora; and, 2) the time commitment necessary to accomplish this bioprospecting and screening. This is the first time a simulation has been used to model the development of a natural product library and this framework is readily scalable to other geographies or therapeutic targets.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

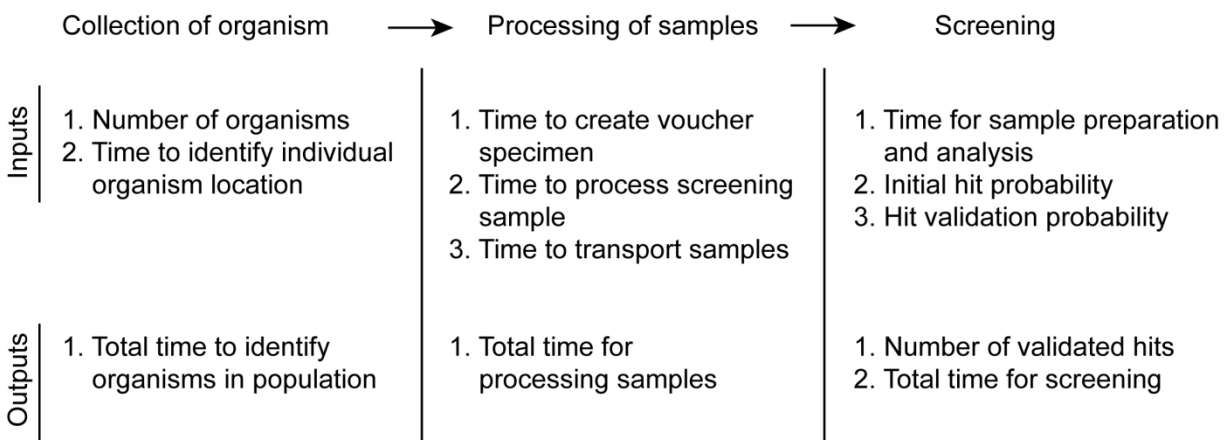


Figure 1. Schematic diagram of the inputs and outputs of the Monte Carlo simulation. These variables capture the primary activities required from identification of the sample to results of the initial screening. These values and their distributions can be dynamically adjusted in the model system.

Methods

A schematic diagram with associated variables was developed describing the process of generating the natural product library of endemic New Zealand terrestrial flora (Figure 1). On the basis of this framework, values for input variables were collected from the literature and a Monte Carlo simulation was developed.

Selection of appropriate model distributions for collection and processing of materials

For the distribution describing the time necessary to find an organism, a Gumbel distribution (Gumbel 2012) was employed.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

$$f(x) = \left(\frac{1}{b}\right) \exp\left(-\frac{x-a}{b}\right) \exp\left[-exp\left(-\frac{x-a}{b}\right)\right]$$

109

110 This distribution allowed modelling the maximum extreme of this variable (∞ is the upper limit,
111 indicating an organism was never found) which results in a conservative model compared to
112 standard modelling used for forecasting species distribution (Araújo & New 2007; Brown 1984)
113 and finding new species (Hill 1979).

114

115 For the distribution describing the time necessary to process samples, create voucher specimens
116 and transport samples, a Weibull (Weibull 1951) distribution was employed:

117

$$f(x) = \alpha\beta^{-\alpha}x^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right)$$

118

119 This distribution is typically used in modelling industrial processes to describe the probability of
120 events such as time to machine failure (Cassady & Kutanoglu 2005; Nelson 1979). Practically,
121 in the current model this distribution is positively skewed indicating there is a probability for the
122 activity to take longer than expected if timing values were distributed normally. The numeric
123 characteristics for these distributions where derived from field experience with these activities
124 (Table 1; (Buenz et al. 2006; Buenz et al. 2007a; Buenz et al. 2007b; Buenz et al. 2007c).

125

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

126 Table 1. Variables and characteristics of the distributions

Variable	Characteristics
Time to find an organism	a=10, b=4, x=10
Time to process	α=2, β=1, x=4
Time to voucher	α=2, β=2, x=9
Time to transport	α=2, β=7, x=25
Validation of hit	μ=0.32, σ=0.24
Time per screening	μ=120, σ=6.2

127

128

129 ***Selection of appropriate model distributions for hit rate and validation of results***

130

131 For the distribution describing the hit rate of antibacterial high-throughput screening, a discrete
132 distribution (Kleywegt et al. 2002) of values from the literature were used. These data were
133 identified from studies that used a terrestrial natural product library and incorporated some type
134 of extraction or processing (Table 2).

135

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

136 Table 2. Potential hit rates for antibacterial screening activities

Hits/1,000	Source of material	Target System	Reference
46	Broad library	Bacterial ribosome	(Lowell et al. 2015)
27	Microbial extracts	Aspergillus fumigatus whole cells	(Monteiro et al. 2012)
27	Secondary metabolites from endophytic fungi	Whole cell assay with multidrug-resistant Pseudomonas aeruginosa	(Zhou et al. 2011)
32.5	Myxobacteria secondary metabolites	Whole-cell assay w/ Vibrio cholera	(Sergeev et al. 2014)
40	Plant extracts (Brazil)	Whole-cell w/ 4 bacteria strains	(Younes et al. 2007)

137

138 For the distribution describing the validation of hits, reported failure rates from the literature
139 32%+/- 24% (Bains 2004; Steinmeyer 2006; Yu 2012) were used and assigned a normal
140 distribution. The distribution describing the time to conduct this analysis were assumed normal
141 using parameters from the clinical laboratory literature to describe the standard deviation (6.3) of
142 a similar laboratory assay (Espy et al. 2006).

143

144 All time data were normalized to full-time equivalent (FTE) effort using 1645 hours per FTE
145 based on a 35 hour work week and with 47 weeks of work per year.

146

147

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

148 ***Modelling system***

149

150 The general calculation structure for the model was built in Excel 2010 (Microsoft, Redmond,
151 WA, USA) and @Risk™ (Palisade, Ithaca, NY, USA) was used as the stochastic modelling.
152 SigmaPlot™ (Systat Software, San Jose, CA, USA) was used for statistical analyses. A total of
153 100,000 iterations were performed using the simulation. The entire model is available as the
154 appendix.

155

156 **Results**

157

158 Figure 2 illustrates the number of projected validated antibacterial hits from five groups of
159 endemic terrestrial New Zealand flora using the stochastic model. The 90% range described
160 below the distribution highlights the numeric values of the upper 5% and lower 5% values
161 determined with respective relative frequency in Figure 2A.

162

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

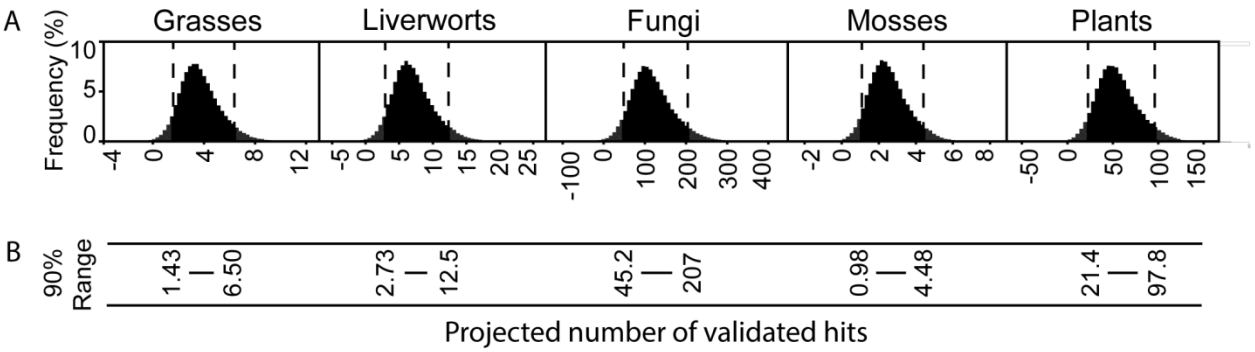
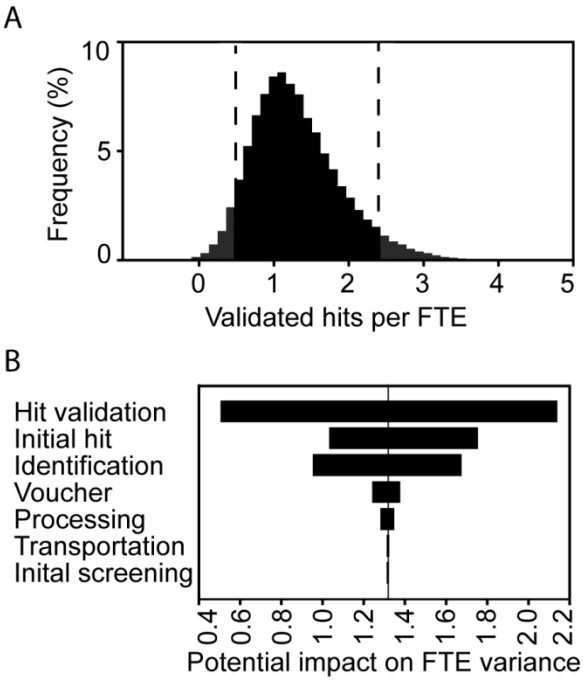


Figure 2. The number of projected validated antibacterial hits from five groups of endemic terrestrial New Zealand flora. Using the Monte Carlo simulation frequency plots where developed (2A). The 90% range described below the distribution highlights the numeric values of the 5% upper and lower values determined by the stochastic model with respective relative frequency (2B).

Figure 3A displays the relative frequency of the validated antibacterial hits per FTE. Importantly, these data suggest that in only 5% of the model iterations less than 0.5 / FTE validated antibacterial leads were identified suggesting that there is little risk of failure to identify a new lead. Additionally, a sensitivity analysis (Figure 3B) shows that the two most significant contributions to the FTE variability for this bioprospecting project are the laboratory activities of screening validating hits and the initial screening activities.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

178



179

180

181 **Figure 3. The number of validated antibacterial hits per full-time equivalent (FTE)**

182 **contribution and variable impact on the outputs.** A sensitivity analysis (3B) shows the

183 contributions to the FTE variability: hit validation = the subsequent analysis to confirm an initial

184 hit in the screening activity; Initial hit = the execution of the initial high-throughput screening

185 activity; Identification = the process of finding and identifying an organism for accession;

186 Voucher = the process of creating a voucher sample for sample confirmation; Processing = the

187 activity of creating the samples necessary for subsequent screening programs; Transportation =

188 moving the samples from the collection to the storage site; and Initial screening = development

189 of the initial screening assay.

190

191

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

192 **Discussion**

193

194 Natural product libraries are an important resource for screening new potential drug leads with
195 unparalleled diversity (Harvey et al. 2015a). While many combinatorial chemistry libraries are
196 large, natural product libraries can be much smaller because of the structural diversity and other
197 favourable characteristics (Zhu et al. 2012). As examples, a library from Brazil has 640 samples
198 (Valli et al. 2013) and a collection from African medical plants has ~1,000 samples (Ntie-Kang
199 et al. 2013) both of which have been identified as valuable tools in drug discovery (Harvey et al.
200 2015b; Szelag et al. 2015).

201

202 Previous smaller-scale studies of New Zealand endemic flora have reported potential
203 antibacterial properties (Earl et al. 2010; Lorimer et al. 1996; Perry & Foster 1995; Ren et al.
204 2014). Furthermore, earlier work screening ~200 New Zealand vascular plants suggested high
205 rates of antibacterial properties (Calder et al. 1986), and one of these leads resulted in
206 identification of a novel antibacterial compound (Hickey et al. 1990). However, these earlier
207 efforts did not use contemporary techniques such as Wyeth fractionation that reduce false
208 positives (Appleton et al. 2007).

209

210 Investigators in Australia have developed the Queensland Compound Library (Frearson & Collie
211 2009; Simpson & Poulsen 2014) which includes >216,000 fractionated natural compounds in the
212 Nature Bank collection and >22,000 extracts in the AIMS Natural Product Collection. New
213 Zealand does not have similar natural product libraries yet the synthetic analysis presented here
214 suggests that endemic New Zealand terrestrial flora can be used to generate an endemic natural

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

product library. As globally 30% of the more promising medicinal plants are threatened with extinction (Brower 2008) and only ~15% of plants have been examined (Saklani & Kutty 2008), there should be a sense of urgency around developing this underappreciated resource.

This study has four limitations. First, data used to develop this model are from disparate sources and thus there is likely systemic bias. For example, the publications used to identify hit rates employed different thresholds for categorization as a hit. While using a discrete distribution in the model allows an averaging effect (Kleywegt et al. 2002), there will be variability in the number of hits based on the threshold selected. Second, this calculation assumes that for each organism only one sample is tested. It is more likely that multiple samples (and furthermore even fractions) would be created, thus potentially increasing the number of antibacterial hits and increase time to screen the entire library (Harvey et al. 2015b). Third, the data around bioprospecting certain groups of organisms for new antibiotics are sparse. For example, there is reason to believe that New Zealand liverworts are a promising resource for antibacterial leads (Lorimer et al. 1996) and thus we may be underestimating the potential hit rate for this group of organisms. Fourth, this analysis only focuses on antibacterial potential and that class of therapeutics may not be the most promising type of therapeutic to pursue in New Zealand terrestrial endemic flora.

The goal of developing this model and performing these analyses was to determine the potential number of novel antibacterial leads that could be identified via high-throughput screening of endemic New Zealand terrestrial flora. Certainly there are novel compounds with antibacterial properties that have already been identified in New Zealand endemic flora (Hickey et al. 1990).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

238 However, without a broad comprehensive screening effort, likely lead candidates in endemic
239 terrestrial flora are potentially overlooked.
240
241 **Conclusions**
242
243 Natural product libraries are valuable tools in the drug discovery process. Here I have shown
244 that developing a natural product library of New Zealand endemic terrestrial flora should contain
245 new, validated hits in a cell-based antibacterial screen. Clearly, once a natural product library is
246 developed, a wide variety of screening paradigms and targets can be evaluated. The high level of
247 endemism in New Zealand makes developing this type of library particularly promising and the
248 success of analogous libraries in other world areas suggests a high-potential for success.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

References

Albuquerque UP, de Medeiros PM, Ramos MA, Júnior WSF, Nascimento ALB, Avilez WMT, and de Melo JG. 2014. Are ethnopharmacological surveys useful for the discovery and development of drugs from medicinal plants? *Revista Brasileira de Farmacognosia* 24:110-115.

Appleton DR, Buss AD, and Butler MS. 2007. A simple method for high-throughput extract prefractionation for biological screening. *CHIMIA International Journal for Chemistry* 61:327-331.

Araújo MB, and New M. 2007. Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22:42-47.

Atanasov AG, Waltenberger B, Pferschy-Wenzig E-M, Linder T, Wawrosch C, Uhrin P, Temml V, Wang L, Schwaiger S, and Heiss EH. 2015. Discovery and resupply of pharmacologically active plant-derived natural products: a review. *Biotechnology Advances*.

Bains W. 2004. Failure rates in drug discovery and development. *Drug Discovery*:9.

Berdy J. 2012. Thoughts and facts about antibiotics: where we are now and where we are heading. *Journal of Antibiotics* 65:385-395. 10.1038/ja.2012.27

Brower V. 2008. Back to nature: extinction of medicinal plants threatens drug discovery. *Journal of the National Cancer Institute* 100:838-839.

Brown JH. 1984. On the relationship between abundance and distribution of species. *American Naturalist*:255-279.

Buenz EJ, Bauer BA, Johnson HE, Tavana G, Beekman EM, Frank KL, and Howe CL. 2006. Searching historical herbal texts for potential new drugs. *BMJ* 333:1314-1315. 10.1136/bmj.39008.492361.BE

Buenz EJ, Bauer BA, Motley TJ, and Limburg PJ. 2007a. Cytotoxic properties of Diospyros seychellarum extract. *Journal of Toxicological Sciences* 32:487-493.

Buenz EJ, Bauer BA, Schnepple DJ, Wahner-Roedler DL, Vandell AG, and Howe CL. 2007b. A randomized Phase I study of Atuna racemosa: a potential new anti-MRSA natural product extract. *Journal of Ethnopharmacology* 114:371-376. 10.1016/j.jep.2007.08.027

Buenz EJ, Tillner JE, Jr., Limburg P, and Bauer BA. 2007c. Antibacterial properties and toxicity of Atuna racemosa extract depend on kernel maturity. *Journal of Ethnopharmacology* 111:592-597. 10.1016/j.jep.2007.01.020

Calder VL, Cole A, and Walker J. 1986. Antibiotic compounds from New Zealand plants. III: a survey of some New Zealand plants for antibiotic substances. *Journal of the Royal Society of New Zealand* 16:169-181.

Cassady CR, and Kutanoglu E. 2005. Integrating preventive maintenance planning and production scheduling for a single machine. *Reliability, IEEE Transactions on* 54:304-309.

Clardy J, Fischbach MA, and Walsh CT. 2006. New antibiotics from bacterial natural products. *Nature Biotechnology* 24:1541-1550.

Cragg GM, and Newman DJ. 2013. Natural products: a continuing source of novel drug leads. *Biochimica et Biophysica Acta (BBA)-General Subjects* 1830:3670-3695.

De Lange P, Norton D, Courtney S, Heenan P, Barkla J, Cameron E, Hitchmough R, and Townsend A. 2009. Threatened and uncommon plants of New Zealand (2008 revision). *New Zealand Journal of Botany* 47:61-96.

De Lange PJ, Sawyer JWD, Rolfe J, and Network NZPC. 2006. *New Zealand indigenous vascular plant checklist*: New Zealand Plant Conservation Network Wellington.

Earl EA, Altaf M, Murikoli RV, Swift S, and O'Toole R. 2010. Native New Zealand plants with inhibitory activity towards Mycobacterium tuberculosis. *BMC Complement Altern Med* 10:25. 10.1186/1472-6882-10-25

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

Espy M, Uhl J, Sloan L, Buckwalter S, Jones M, Vetter E, Yao J, Wengenack N, Rosenblatt J, and Cockerill F. 2006. Real-time PCR in clinical microbiology: applications for routine laboratory testing. *Clinical Microbiology Reviews* 19:165-256.

Fischbach MA, and Walsh CT. 2009. Antibiotics for emerging pathogens. *Science* 325:1089-1093.

Frearson JA, and Collie IT. 2009. HTS and hit finding in academia–from chemical genomics to drug discovery. *Drug Discovery Today* 14:1150-1158.

Gu J, Gui Y, Chen L, Yuan G, Lu H-Z, and Xu X. 2013. Use of natural products as chemical library for drug discovery and network pharmacology.

Gu R, Wang Y, Long B, Kennelly E, Wu S, Liu B, Li P, and Long C. 2014. Prospecting for bioactive constituents from traditional medicinal plants through ethnobotanical approaches. *Biological and Pharmaceutical Bulletin* 37:903-915.

Gumbel EJ. 2012. *Statistics of extremes*: Courier Corporation.

Gyllenhaal C, Kadushin MR, Southavong B, Sydara K, Bouamanivong S, Xaiveu M, Xuan LT, Hiep NT, Hung NV, Loc PK, Dac LX, Bich TQ, Cuong NM, Ly HM, Zhang HJ, Franzblau SG, Xie H, Riley MC, Elkington BG, Nguyen HT, Waller DP, Ma CY, Tamez P, Tan GT, Pezzuto JM, and Soejarto DD. 2012. Ethnobotanical approach versus random approach in the search for new bioactive compounds: support of a hypothesis. *Pharmaceutical Biology* 50:30-41. 10.3109/13880209.2011.634424

Harvey AL. 2008. Natural products in drug discovery. *Drug Discovery Today* 13:894-901.

Harvey AL, Edrada-Ebel R, and Quinn RJ. 2015a. The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery* 14:111-129.

Harvey AL, Edrada-Ebel R, and Quinn RJ. 2015b. The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews: Drug Discovery* 14:111-129. 10.1038/nrd4510

Hickey B, Lumsden A, Cole A, and Walker J. 1990. Antibiotic compounds from New Zealand plants: methyl haematommate, an anti-fungal agent from Stereocaulon ramulosum. *New Zealand Nat Sci* 17:49-53.

Hill BM. 1979. Posterior moments of the number of species in a finite population and the posterior probability of finding a new species. *Journal of the American Statistical Association* 74:668-673.

Kleywegt AJ, Shapiro A, and Homem-de-Mello T. 2002. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12:479-502.

Li JW, and Vederas JC. 2009. Drug discovery and natural products: end of an era or an endless frontier? *Science* 325:161-165. 10.1126/science.1168243

Lipinski C, and Hopkins A. 2004. Navigating chemical space for biology and medicine. *Nature* 432:855-861. 10.1038/nature03193

Lorimer S, Barns G, Evans A, Foster L, May B, Perry N, and Tangney R. 1996. Cytotoxicity and antimicrobial activity of plants from New Zealand's subantarctic islands. *Phytomedicine* 2:317-323.

McGlone M, Duncan R, and Heenan P. 2001. Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand. *Journal of Biogeography* 28:199-216.

Michelson S, Sehgal A, and Friedrich C. 2006. In silico prediction of clinical efficacy. *Curr Opin Biotechnol* 17:666-670. 10.1016/j.copbio.2006.09.004

Mortimer N. 2004. New Zealand's geological foundations. *Gondwana Research* 7:261-272.

Nelson PR. 1979. Control charts for Weibull processes with standards given. *Reliability, IEEE Transactions on* 28:283-288.

Ntie-Kang F, Zofou D, Babiaka SB, Meudom R, Scharfe M, Lifongo LL, Mbah JA, Mbaze LMa, Sippl W, and Efange SM. 2013. AfroDb: a select highly potent and diverse natural product library from African medicinal plants.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017

Perry NB, and Foster LM. 1995. Sesquiterpene/quinol from a New Zealand liverwort, Riccardia crassa. *Journal of natural products* 58:1131-1135.

Ren L, Hemar Y, Perera CO, Lewis G, Krissansen GW, and Buchanan PK. 2014. Antibacterial and antioxidant activities of aqueous extracts of eight edible mushrooms. *Bioactive Carbohydrates and Dietary Fibre* 3:41-51.

Saklani A, and Kutty SK. 2008. Plant-derived compounds in clinical trials. *Drug discovery today* 13:161-171.

Schwikkard SL, and Mulholland DA. 2014. Useful methods for targeted plant selection in the discovery of potential new drug candidates. *Planta Medica* 80:1154-1160.

Simpson M, and Poulsen S-A. 2014. An overview of Australia's compound management facility: the Queensland Compound Library. *ACS chemical biology* 9:28-33.

Steinmeyer A. 2006. The Hit-to-Lead Process at Schering AG: Strategic Aspects. *ChemMedChem* 1:31-36.

Sucher NJ. 2013. The application of Chinese medicine to novel drug discovery. *Expert opinion on drug discovery* 8:21-34.

Szelag M, Czerwoniec A, Wesoly J, and Bluyssen H. 2015. Identification of STAT1 and STAT3 Specific Inhibitors Using Comparative Virtual Screening and Docking Validation. *PLoS One* 10:e0116688.

Tao L, Zhu F, Qin C, Zhang C, Xu F, Tan CY, Jiang YY, and Chen YZ. 2014. Nature's contribution to today's pharmacopeia. *Nature Biotechnology* 32:979-980. 10.1038/nbt.3034

Valli M, Dos Santos RN, Figueira LD, Nakajima CH, Castro-Gamboa I, Andricopulo AD, and Bolzani VS. 2013. Development of a natural products database from the biodiversity of Brazil. *Journal of natural products* 76:439-444.

Weibull W. 1951. A statistical distribution function of wide applicability. *Journal of applied mechanics* 103.

Yu MJ. 2012. Simulating the drug discovery pipeline: a Monte Carlo approach. *J Cheminform* 4:32. 10.1186/1758-2946-4-32

Zhu F, Ma XH, Qin C, Tao L, Liu X, Shi Z, Zhang CL, Tan CY, Chen YZ, and Jiang YY. 2012. Drug discovery prospect from untapped species: indications from approved natural product drugs. *PLoS One* 7:e39782. 10.1371/journal.pone.0039782

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2734v1 | CC BY 4.0 Open Access | rec: 18 Jan 2017, publ: 18 Jan 2017