Marco van Zwetselaar

Bioinformatician / Software Developer

zwets@kcri.ac.tz · io.zwets.it · +255 782 334 124

P.O. Box 2236 · Moshi, Kilimanjaro · Tanzania

Bioinformatician with research focus on whole genome sequencing and microbial genomics.
Thirty years work experience as an IT professional. Background in modelling and statistics.
Open source software developer since the appearance of the personal computer.

CURRENT

Bioinformatician at Kilimanjaro Clinical Research Institute (KCMC, Moshi, Tanzania), Aug 2012 – present

- construct the analysis pipelines of the genomics unit at KCRI, and design and perform bespoke analyses of genomic data for various projects
- server management, data management, software development, analysis & visualisation for clinical and genomic studies at KCRI

PREVIOUS

Enterprise architect: Atos, Aug 2005 – Aug 2012

Lead architect: Delta Lloyd OHRA (through Atos), 2011 – 2012

Consultant: a.s.r. verzekeringen (through Atos), 2008 – 2011

IT architect: Achmea (through Atos Origin), 2005 – 2009

IT architect: Fortis, 1998 – 2004

Software architect: RIVM, 1994 – 1998

Scientific programmer: Utrecht University, 1990 – 1994

LANG-UAGES

Spoken: English, Dutch, Italian, German, French, some Swahili

Computing: C / C++, Java, Haskell, Python, R, awk/bash, JavaScript, Scheme, Erlang, Smalltalk, Prolog

---

RECENT COURSES

Mathematics for Machine Learning, Imperial College (Coursera), July 2020, licence 4CHZR56VKCFG

Real-time Virus Genome Sequencing, workshop on Nanopore sequencing for viral outbreak response, ARTIC Network, Ghana, Dec 2018

Specialization Probabilistic Graphical Models (Honours), Stanford (Coursera) Mar 2017, licence 6KVJTGM283MK

Specialization Bioinformatics Algorithms (Distinction), UCSD (Coursera) Apr 2016, licence UQQK6D5JX8DV

Classical Papers in Molecular Genetics, University of Geneva (Coursera) Sep 2016, licence Y8WYEP2KQH2P.

High-dimensional Data Analysis, PH525.4.x (edX), Apr 2016, licence fd79ffa5aaed4e829401a42a1c548b9d

Statistical Inference for High-throughput Experiments, PH525.3.x (edX), Mar 2016, licence aec1b523e2d2418cbc0ab47d8ab5b35b

Algorithms for DNA sequencing, Johns Hopkins (Coursera), Jan 2016, licence GMWNE2NMP8

Extremes of Life: Microbes and their Diversity, Kyoto University (Coursera) Dec 2014, licence 2495a8d27d25416996af852e816cd78b

---

RECENT PUBLIC-ATIONS

Nkya S, Mwita L, Mgaya J, Kumburu H, van Zwetselaar M, Menzel S, et al. Identifying genetic variants and pathways associated with extreme levels of fetal hemoglobin in sickle cell disease in Tanzania. BMC Med Genet. 2020;21:125.

Kumburu HH, Sonda T, van Zwetselaar M, *et al.* Using WGS to identify antibiotic resistance genes and predict antimicrobial resistance phenotypes in MDR Acinetobacter baumannii in Tanzania. *J Antimicrob Chemother.* 2019 doi:10.1093/jac/dkz055.4

Sonda T, Kumburu H, Zwetselaar M van, *et al.* Molecular epidemiology of virulence and antimicrobial resistance determinants in Klebsiella pneumoniae from hospitalised patients in Kilimanjaro, Tanzania. *Eur J Clin Microbiol Infect Dis* 2018; 1–14

Sonda T, Kumburu H, van Zwetselaar M, *et al.* Whole genome sequencing reveals high clonal diversity of Escherichia coli isolated from patients in a tertiary care hospital in Moshi, Tanzania. *Antimicrobial Resistance & Infection Control* 2018; 7.

Van Zwetselaar M, Nyombi B, Sonda T, *et al.* Aeromonas caviae mimicking Vibrio cholerae infectious enteropathy in a cholera-endemic region with possible public health consequences: two case reports. *Journal of Medical Case Reports* 2018; 12.

Sonda T, Kumburu H, Zwetselaar M van, *et al.* Prevalence and risk factors for CTX-M gram-negative bacteria in hospitalized patients at a tertiary care hospital in Kilimanjaro, Tanzania. *Eur J Clin Microbiol Infect Dis* 2018; 1–10.

Ostermann J, Vasudevan L, Van Zwetselaar M, Moses S, Engadaya E, Mfinanga S. Mobile Phone Assisted Reminder and Incentive System (mParis). *NIH Tech Showcase presentation 2018.*

RECENT SOFT-WARE

github.com/zwets | io.zwets.it

- **unfasta** – command-line pipes and filters for genomic data
- **taxo** – offline command-line taxonomy browser
- **kcst** – k-mer counting sequence typing and species identification
- **sast** – simple assembly sequence typing
- **fastq-utils** – small utilities to manipulate fastq format files
- **gene-cutter** – template-based extraction of allele matches
- **blast-pcr** – test primers against a BLAST database