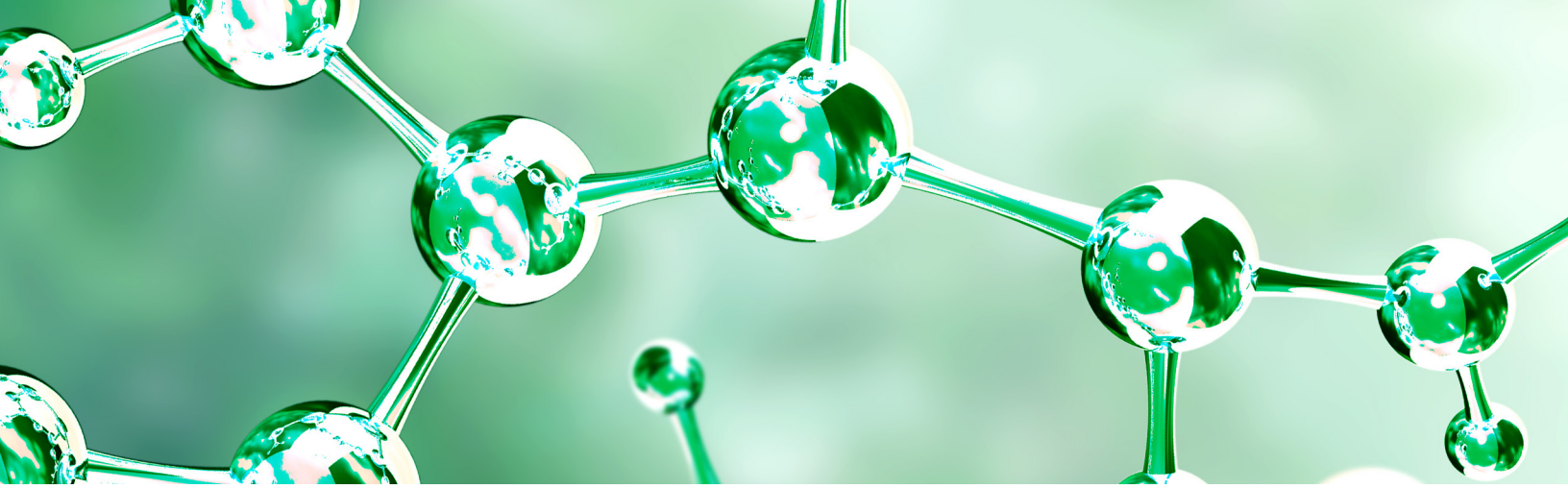
ALTASCIENCES

WHOLE GENOME SEQUENCING, PROTEOMICS, AND FUNCTION CHARACTERIZATION OF THE SINCLAIR NANOPIG™ FOR (BIO) PHARMACEUTICALS SAFETY ASSESSMENT

Yafei Chen, Wendell P. Davis
Altasciences



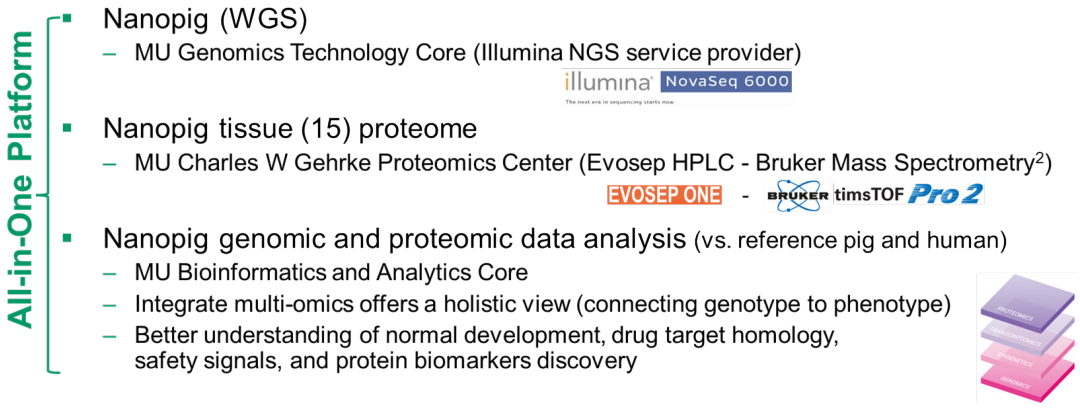
Introduction

Safety and efficacy of human (bio)pharmaceuticals rely heavily on their interaction with therapeutic targets. The Sinclair minipig has represented a valuable species for biomedical research for decades, but the lack of knowledge about the human orthologous targets has hampered its use in preclinical studies, especially with those incorporating biologics. The current whole genome sequencing (WGS) and tissue-based proteomics, combined with advanced bioinformatics technologies, have enabled the closing of the translational knowledge gap (minipig versus human), and the development of new digital research tools to improve the sustainability and quality of biomedical research and development.

The Sinclair Nanopig™ was recently introduced by Altasciences and Sinclair Bio Resources' Nanopig breeding and colony facility as the next-generation non-rodent model for drug safety assessment. We have generated and reported reference values for the Sinclair Nanopig™, which include a limited growth rate and lower body weight. Our findings indicate similar clinical pathology data, organ weights, and background microscopic findings compared to other minipig breeds (Chen et al., 2023). Genome-based comparison of drug targets with quantitative tissue protein expression analysis allows rational prediction of pharmacology, cross-reactivity, and potential toxicity of human drugs in animal models—improving clinical translation and drug attrition (Vamathevan et al., 2013; Heckel et al., 2015). The white paper provides genomic, proteomic, and functional characterization data—covering metabolism and immune systems—of the Sinclair Nanopig™. These scientific justifications support the selection of this human-relevant animal species for regulatory pharmacology and drug safety assessment. Additionally, this data expands translational knowledge of the Sinclair Nanopig™, helping to reduce and replace traditional non-rodent models in drug development.

The Sinclair Nanopig™ characterization project was led by Altasciences in close collaboration with academic and industry partners, including Sinclair Bio Resources, the University of Missouri (Figure 1), and BioIVT (CYP450 enzymatic activity). With the sequencing data of the Sinclair Nanopig™ genome and proteome (Chen et al., 2024), researchers are now able to design cross-species assessment of pharmacology-related adverse effects, enabling deeper understanding of drug metabolism and off-target-related toxicities.

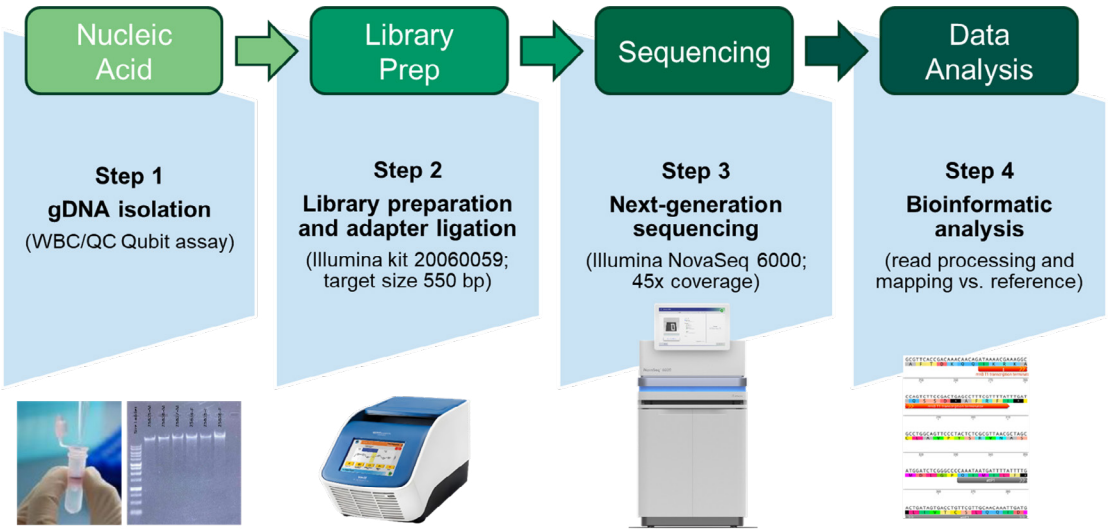
Figure 1. Altasciences and the University of Missouri (MU) Collaboration



Whole Genome Sequencing of the Sinclair Nanopig™

A simplified workflow diagram for a four-step process of the Sinclair Nanopig™ WGS is shown in Figure 2. Briefly, six Nanopigs (3 males and 3 females; four months old) from Sinclair Bio Resources were used to isolate genomic DNA (Qiagen DNeasy Blood and Tissue Kit #69504) from circulating white blood cells (WBC) at Altasciences' Preclinical Columbia facility, followed by gDNA quantification (Invitrogen Qubit #Q33230) and quality control by integrity assessment (100 ng on 1% agarose gel) for subsequent library preparation (Illumina kit #20060059; 500 ng gDNA; target insert size 550 bp) at the University of Missouri Genomic Technology Core. All six library samples (8.75-13.2 ng/µL; fragment size 691-745 bp) were sequenced on NovaSeq 6000 PE150 flowcell according to Illumina's standard next generation sequencing (NGS) protocol with a 45-fold depth of coverage. The Illumina paired-end sequencing methods were applied, and clean reads were mapped to the reference Duroc pig (NCBI sscrofa11.1; Ensembl Size 2,501,912,388 bp) to obtain a Nanopig genome assembly conducted by the University of Missouri Bioinformatics and Analytics Core.

Figure 2. Nanopig whole genome sequencing workflow

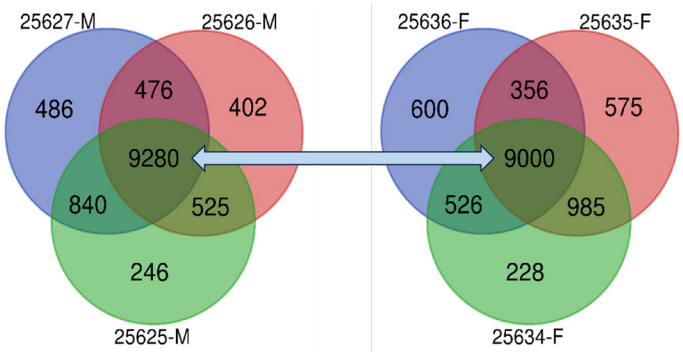


For the first time, a draft Nanopig genome was assembled at full chromosome-scale resolution with a total length of 2.9 Gb. The alignment of clean sequencing reads to the reference Duroc pig genome revealed 98-99% of total mapped reads (Figure 3a). Variant annotation indicates that the replicates of both the male (M) and female (F) Nanopigs are very consistent, and have high repeatability (Figure 3b).

Figure 3a. Mapping coverage of genome sequences (Nanopigs vs. Duroc pig)

No.	Sample	Total mapped
1	25625-M	825,864,920 (98.47%)
2	25626-M	851,289,469 (98.47%)
3	25627-M	975,456,368 (98.36%)
4	25634-F	866,940,025 (98.24%)
5	25635-F	856,232,831 (98.24%)
6	25636-F	882,552,801 (98.42%)

Figure 3b. Variant annotation of Nanopigs genome sequences (male vs. female)



The annotation analysis of Sinclair Nanopig™ genomic sequences with reference to the genome of Duroc pig revealed that Nanopig WGS data was obtained in high-quality. The Nanopig genome assembled at chromosome level with a total length of 2.9 Gb versus the Duroc pig with a total of 2.5 Gb, and revealed alignment and mapping coverage of sequences >98% with no clear and substantial genomic variance. Importantly, variant annotation indicates male and female Nanopig replicates are consistent and have high repeatability with >9,000 genes overlapped. These results indicate distinct characteristics of Nanopigs derive from small-scale alterations in the genome (single nucleotide polymorphisms or translational modifications), rather than large-scale deletion or insertion polymorphisms. Furthermore, genome annotation was performed to determine the similarities and differences between Nanopigs and human (GRCh38.p14; Ensembl Size 3,099,750,718 bp) database underlying metabolism and immune systems that are more relevant to the drug or biologics safety assessment. The Nanopig genome with a total length of 2.9 Gb is similar to the human genome of 3.1 Gb. To further elucidate the genetic basis of Nanopig metabolism and immune systems, human comparative genomic analyses (gene homology) revealed 11,711 metabolism-related genes and 1,606 immunity-related genes overlapped with the Nanopig.

The cytochrome P450 enzymes (CYPs), mainly found within the mitochondria and endoplasmic reticulum of liver cells, are crucial and responsible for the majority of drug metabolism in humans [Zanger, et al., 2013]. Specifically, 57 known human genes encode for the different CYP enzymes, and these have been classified into 18 protein families [Goh et al., 2017]. The enzymes in the first three families (CYP1, CYP2, and CYP3) play a pivotal role and are responsible for the metabolism of 70–80% of clinical drugs [Ingelman-Sundberg 2004; Lynch et al., 2007]. Within these families, enzymes CYP1A1, CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C18, CYP2C19, CYP2D6, CYP2E1, CYP3A4, and CYP3A5 notably metabolize most of the clinical drugs [Zanger, et al., 2013; Zhao, et al., 2021]. In the Nanopig, a total of 47 CYP450 genes were identified, with 20 in the CYP450 enzyme families 1, 2, and 3 (including CYP1A, CYP2A, 2C, 2D, 2E and CYP3A) that are very similar to human 57 CYP450 genes with 24 in CYP450 1, 2, and 3 families. However, factors that affect CYP expression with diverse enzymatic activity have a major impact on drug safety and efficacy outcomes. Therefore, a subsequent proteomics approach was used to investigate the genetic links between metabolism-related genes, as well as innate and adaptive immunity-related genes, and tissue protein expression abundance.

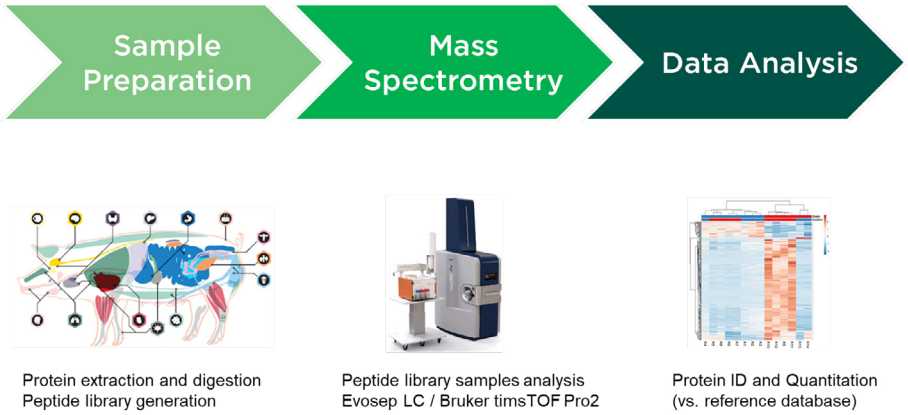
Proteomics of the Sinclair Nanopig™

To uncover genetic links to tissue protein abundance, an advanced liquid chromatography with tandem mass spectrometry (LC-MS-MS) proteomics platform was used to characterize protein expression profiles in 15 selected Nanopig tissues with pharmaceutical relevance, including:

1. Plasma
2. Brain (front lobe cortex)
3. Adrenal gland
4. Lung
5. Liver (left lobe)
6. Kidney (cortex)
7. Stomach
8. Small intestine (duodenum)
9. Colon
10. Testis or Ovary
11. Skeletal muscle
12. Heart (left ventricle)
13. Spleen
14. Thymus
15. Lymph nodes (cervical)

The Nanopigs euthanasia and tissue collection was conducted by Altasciences' histology team at the Columbia site according to the facility's standard operating procedures. All harvested tissue samples were stored at -70-80 °C until processed and analyzed at the University of Missouri Proteomics Center (Figure 4).

Figure 4. Nanopig tissue-based proteomics workflow

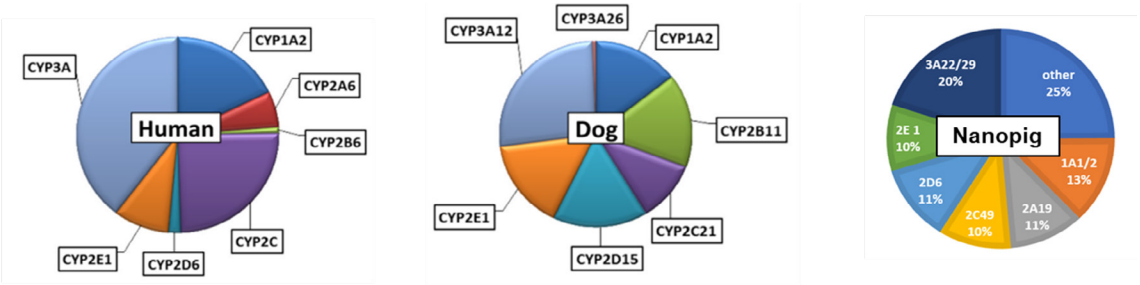


Each tissue sample was briefly ground in a denaturing buffer and spun at 16,000xg for 20 minutes to extract proteins, followed by alkylation and precipitation with cold acetone at -20 °C overnight. Protein pellets were redissolved in urea and protein concentration was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific). Twenty micrograms (µg) of protein were digested with proteinase Lys-C and trypsin in urea buffer. To construct a comprehensive spectral library, 15 µg of digested peptides from each tissue sample were fractionated using a high pH reversed-phase spin column with a peptide fractionation kit (Thermo Fisher Scientific) resulting in the acquisition of 8 fractions for each tissue. Following fractionation, 500 ng of the digested peptide from each sample was acquired with the data-independent acquisition combined with parallel accumulation serial fragmentation (DIA-PASEF) method. For subsequent sequencing data acquisition, an EvoSep One (Denmark) liquid chromatography system was utilized and coupled online to a modified trapped ion mobility spectrometry quadrupole time-of-flight mass spectrometer (timsTOF Pro 2, Bruker Daltonik, GmbH, Germany) via a nanoelectrospray ion source (CaptiveSpray). Spectronaut® (version 18.3) was used for constructing the spectral library and analyzing the obtained DIA-PASEF data. The Spectronaut output was analyzed using the MSstats package (version 4.6.3) within the R environment (version 4.3.1) to ascertain protein abundance, differential expression fold changes, and associated p-values.

The Sinclair Nanopig™ proteome analysis was performed to determine the similarities and differences (human orthologous) between the reference human (Uniprot-UP000005640; 20,381 protein counts) protein sequence databases, focusing on drug metabolism enzymes and transporters (DMET) and immune (innate and adaptive responses) systems due to their essential role in drug disposition and biologics immunosafety assessment. First, the Nanopig protein expression profile was successfully achieved in 14 major tissue and plasma samples. The human orthologous of Phase I/II enzymes and transporters (influx/efflux) were further identified in three relevant Nanopig tissues (for drug disposition), including the liver (44 Phase I and 27 Phase II enzymes; 96 transporters), small intestine (18 Phase I and 17 Phase II enzymes; 116 transporters), and kidney (33 Phase I and 21 Phase II enzymes; 152 transporters).

Accordingly, the key CYP450 enzymes quantified in Nanopig liver included 21 isoforms with eight in the CYP450 families 1, 2, and 3 (relative quantities to the total amount of CYP450: 3% CYP1A1, 10% CYP1A2, 11% CYP2A19, 10% CYP2C49, 11% CYP2D6, 10% CYP2E1, 11% CYP3A22, and 9% CYP3A29). Importantly, a relative similarity in the total contents of CYP450 1, 2, and 3 enzymes was observed in humans and beagle dogs (Figure 5). As there are many similarities in metabolic enzymes and transporters to those in humans, the Nanopig is a relevant and powerful animal model for candidate drug in toxicokinetic and regulatory toxicology studies. Next, the correlation analysis between CYP450 protein abundance and enzymatic activity was investigated.

Figure 5. Nanopig liver key CYP450 isoforms abundance (vs. human and beagle dog)



	Human		%		Beagle		%		Nanopig		%
CYP1	1A2	14	14		1A2	14	14		1A1	3	13
									1A2	10	
CYP2	2A6	8	51				58		2A19	11	42
	2B6	2			2B11	16					
	2C	25			2C21	11			2C49	10	
	2D6	3			2D15	16			2D6	11	
	2E 1	13			2E 1	16			2E 1	10	
CYP3	3A	35	35		3A12	27	28		3A22	11	19
					3A26	1			3A29	9	
Court MH. Canine cytochrome P-450 pharmacogenetics. 2013.									CYP4/20/27		25

In addition, the Nanopig immune system includes human orthologous components and effectors (both innate and adaptive), which were profiled and identified in the circulatory system and major lymphoid organs (thymus, spleen, and lymph node) for biologics testing, including:

Plasma	30 antimicrobial proteins, 11 acute phase proteins, 25 cytokines, 3 complements (C3, C5, C8), and immunoglobulin M
Thymus	78 antimicrobial proteins, 54 acute phase proteins, 46 cytokines, 47 T cell receptor and 35 B cell receptor signaling proteins
Spleen	92 antimicrobial proteins, 51 acute phase proteins, 51 cytokines, 50 T cell receptor and 40 B cell receptor signaling proteins
Lymph node	94 antimicrobial proteins, 52 acute phase proteins, 60 cytokines, 55 T cell receptor and 45 B cell receptor signaling proteins

Importantly, the swine leukocyte antigens (SLAs) identified in each Nanopig are critical for immunological reactions. Beyond these humoral and cellular immunity components, further assessment of the utility of Nanopig for immunosafety testing using reference biologics (tested preclinically and clinically) is warranted.

These proteomic data are critical because many cellular mechanisms depend on post-translational modification of proteins and specific drug-protein interactions, especially in drug metabolism and biologics-induced immunogenicity and/or immunotoxicity.

Key Hepatic CYP450 Enzymes Activity in the Sinclair Nanopig™

Correlation analysis between the Nanopig hepatic CYP450 protein abundance and enzymatic activity (K_m and V_{max}) was investigated with the most commonly used probe substrates (*FDA guideline: Drug Development and Drug Interactions | Table of Substrates, Inhibitors and Inducers*) at BioIVT (Kansas City, KS, USA). The experimental conditions and the Nanopig CYP450 enzymes' kinetic constants (K_m and V_{max}) are presented and compared with the data of humans and beagle dogs in Table 1. The K_m values for the CYP2C substrate diclofenac and the CYP3A substrate midazolam were most similar between the Nanopig and humans, with the K_m of CYP2D most similar between the Nanopig and canine and K_m of CYP1A most similar between human and canine. This data should be carefully considered and integrated with the *in vitro* CYP450 assay to select the relevant species for preclinical toxicity studies.

Table 1. Experimental conditions and determination of kinetic constants (K_m and V_{max})

Probe Substrate	Enzyme Subfamily	Final Protein Conc. (mg/mL)	Incubation Time (min)	Concentration Range (μM; 13 concentrations)
Phenacetin	CYP1A	0.1	10	1, 2, 4, 7, 9, 11, 14, 17, 22, 28, 55, 83, 110
Diclofenac	CYP2C	0.1	40	0.5, 1, 2, 5, 10, 15, 20, 25, 31, 38, 50, 63, 125
Dextromethorphan	CYP2D	0.01	20	0.1, 0.2, 0.4, 0.6, 0.8, 1, 1.25, 1.5, 2, 2.5, 5, 7.5, 10
Midazolam	CYP3A	0.1	10	0.2, 0.4, 0.8, 1.2, 1.6, 2, 2.5, 3, 4, 5, 10, 15, 20

Probe Substrate	Enzyme Subfamily	Human K_m	Human V_{max}	Nanopig™ K_m	Nanopig™ V_{max}	Canine K_m	Canine V_{max}
Phenacetin	CYP1A	94.5 ± 3.8	662 ± 13	4.23 ± 0.56	988 ± 45	54 ± 30	790 ± 250
Diclofenac	CYP2C	9.16 ± 0.33	2610 ± 30	7.89 ± 1.15	91.4 ± 5.3	29 ± 8	990 ± 30
Dextromethorphan	CYP2D	11.0 ± 0.9	225 ± 6	0.811 ± 0.070	879 ± 23	0.72 ± 0.16	2600 ± 100
Midazolam	CYP3A	2.52 ± 0.19	807 ± 18	4.06 ± 0.41	621 ± 39	1.5 ± 0.5	270 ± 50

Note: BioIVT study report reference XT234148; K_m: μM; V_{max}: pmol/min/nmol total P450 or mg protein; Mean ± SD

Perspectives of Using the Sinclair Nanopig™ in Drug and Biologics Safety Assessment

This is the first report of a newly revealed chromosome-level-based version of the Sinclair Nanopig™ genome together with a comparative proteomic analysis of tissues with pharmaceutical relevance and critical information about metabolism, and the immune system as the basis for translational research. Functional genomics and proteomics play a crucial role in developing and characterizing Nanopigs for drug and biologics development. They provide deeper insights into the intricate molecular mechanisms governing biopathology and physiology, essential for ensuring the safety and efficacy of drugs and biologics.

With this expansion of translational knowledge of the Sinclair Nanopig™, we enable their increased use and acceptance in the industry, and reduce the use of other traditional non-rodent models in drug development. Our searchable multi-omics database encourage the broad use of the Nanopig for pharmacology, biomarkers discovery, and drug safety assessment.

References

1. Y. Chen, A. Martin, Z. Patenaude, D. Brocksmith, and W. Davis. 2023. "Downsized Sinclair vs. Göttingen Minipigs: Similarities and Differences of Toxicological Reference Range Data in Preclinical Safety Studies." *The Toxicologist*, a Supplement to *Toxicological Sciences*, Abstract #3052.
2. Vamathevan JJ, Hall MD, Hasan S, Woollard PM, Xu M, Yang Y, Li X, Wang X, Kenny S, Brown JR, Huxley-Jones J, Lyon J, Haselden J, Min J, Sanseau P. Minipig and beagle animal model genomes aid species selection in pharmaceutical discovery and development. Toxicol Appl Pharmacol. 2013 Jul 15;270(2):149-57.
3. Heckel T, Schmucki R, Berrera M, Ringshandl S, Badi L, Steiner G, Ravon M, Küng E, Kuhn B, Kratochwil NA, Schmitt G, Kiialainen A, Nowaczyk C, Daff H, Khan AP, Lekolool I, Pelle R, Okoth E, Bishop R, Daubenberger C, Ebeling M, Certa U. Functional analysis and transcriptional output of the Göttingen minipig genome. BMC Genomics. 2015 Nov 14;16:932.
4. Y. Chen, N. Bivens, H. An, B. Mooney, T. Nguyen, L. Coghill, M. Sahoo, T. Madsen, N. Zuidhof, J. Horkman, L. Haupt, M. Evans, R. Campbell, I. Vanterpool, S. Mason, and W. Davis. 2024. "Sinclair Nanopig™: From Multi-Omics Characterization to Pharmacology and Toxicology Validation Underline Drug Metabolism and Immune System." *The Toxicologist: Late-Breaking Supplement*, a Supplement to *Toxicological Sciences*, Abstract #5002.
5. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol Ther. 2013 Apr;138(1):103-41.
6. Goh LL, Lim CW, Sim WC, Toh LX, Leong KP. Analysis of Genetic Variation in CYP450 Genes for Clinical Implementation. PLoS One. 2017 Jan 3;12(1):e0169233. doi: 10.1371/journal.pone.0169233. PMID: 28046094; PMCID: PMC5207784.
7. Ingelman-Sundberg M. Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. Naunyn Schmiedebergs Arch Pharmacol. 2004 Jan;369(1):89-104. doi: 10.1007/s00210-003-0819-z. Epub 2003 Oct 22. PMID: 14574440.
8. Lynch T, Price A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. Am Fam Physician. 2007 Aug 1;76(3):391-6. PMID: 17708140.
9. Zhao M, Ma J, Li M, Zhang Y, Jiang B, Zhao X, Huai C, Shen L, Zhang N, He L, Qin S. Cytochrome P450 Enzymes and Drug Metabolism in Humans. Int J Mol Sci. 2021 Nov 26;22(23):12808. doi: 10.3390/ijms222312808. PMID: 34884615; PMCID: PMC8657965.
10. FDA Guideline: Drug Development and Drug Interactions | Table of Substrates, Inhibitors and Inducers. https://www.fda.gov/drugs/drug-interactions-labeling/drug-development-and-drug-interactions-table-substrates-inhibitors-and-inducers
11. Bendixen E, et al. Advances in porcine genomics and proteomics--a toolbox for developing the pig as a model organism for molecular biomedical research. Brief Funct Genomics. 2010 May;9(3):208-19.
12. Helke KL, et al. Pigs in Toxicology: Breed Differences in Metabolism and Background Findings. Toxicol Pathol. 2016 Jun;44(4):575-90.
13. Schelstraete W, et al. Characterization of Porcine Hepatic and Intestinal Drug Metabolizing CYP450: Comparison with Human Orthologues from A Quantitative, Activity and Selectivity Perspective. Sci Rep. 2019 Jun 25;9(1):9233.
14. Buyssens L, et al. Hepatic Cytochrome P450 Abundance and Activity in the Developing and Adult Göttingen Minipig: Pivotal Data for PBPK Modeling. Front Pharmacol. 2021 Apr 15;12:665644.
15. Arbitrio M, et al. DMET™ (Drug Metabolism Enzymes and Transporters): a pharmacogenomic platform for precision medicine. Oncotarget. 2016 Aug 16;7(33):54028-54050.
16. Rubic-Schneider T, et al. Minipigs in Translational Immunosafety Sciences: A Perspective. Toxicol Pathol. 2016 Apr;44(3):315-24.