STUDY REPORT

Table of Contents

科佰生物Cobioer

PROJECT NUMBER: E6925-T2501

STUDY TITLE: NGS-QC-PANEL test

SPONSOR: Cobioer

SPONSOR REPRESENTATIVE: -

SPONSOR ADDRESS: -

SPONSOR TELEPHONE: -

SPONSOR EMAIL: -

STUDY DIRECTOR: Zhenzhen Song

TELEPHONE: -

EMAIL: zhenzhen.song@crownbio.com

TESTING FACILITY: -

STUDY INITIATION DATE: -

STUDY COMPLETION DATE: -

DATE OF DOCUMENT: -

DOCUMENT AUTHOR: Wubin Qian

VERSION OF DOCUMENT: 1.0

科佰生物Cobioer

OVERVIEW

SPECIMEN DESCRIPTION

Number of Specimens: 1

SampleID	Species
PC-3-20260211	Human

SERVICES PERFORMED

- **Service:** Crown NGS-QC for human/mouse samples
- **Service items:**
    - Sample identification by matching to reference cell line database
    - Human-mouse interspecies contamination quantification
    - Sample genetic heterogeneity quantification
    - Sample contamination check and contaminant identification (if the contaminant is already in the reference cell line database)
    - Mouse syngeneic model identification (if applicable)
    - Mouse strain (if applicable)
    - Microorganism infection detection
    - Mycoplasma contamination detection
    - Gender identification for human samples
    - Genetic admixture analysis for human samples

RESULT SUMMARY

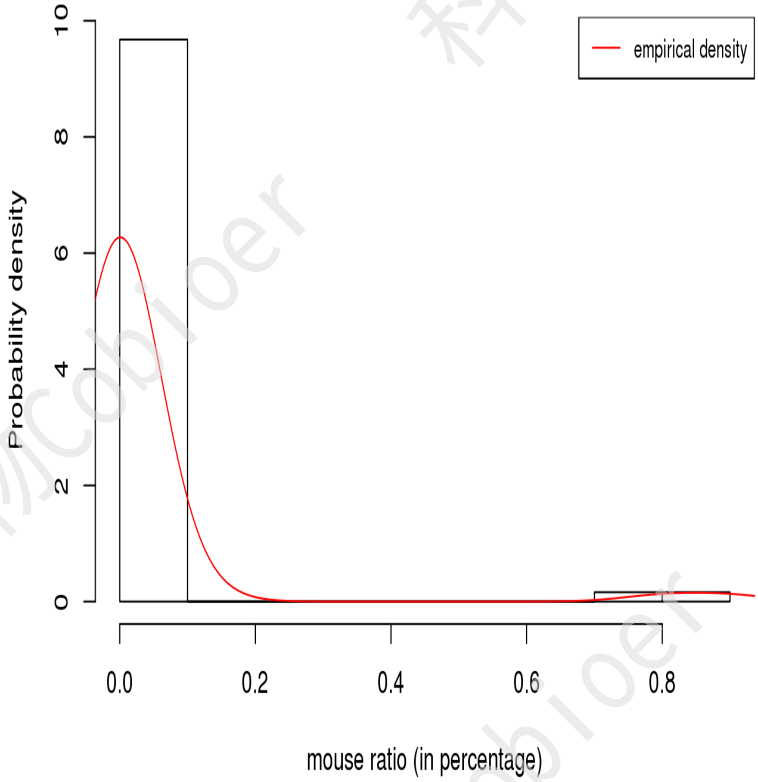
QC Item	Results
Mouse ratio	0%
Matched sample(by SNP)	PC3
Matched sample(by STR)	–
Heterogeneity ratio*	low
Contaminant detected	NA
Mouse model	NA
Microorganism infection	negative
Mycoplasma/host ratio(substrain)	0(-)
Gender (human sample only)	Female
Genetic admixture (CEU: CHB: YRI) **	– : – : –

*: See definition in Reference 1.

**: The three reference populations are: Han Chinese (CHB), Nigeria Yoruba (YRI), and Utah residents with Northern and Western European ancestry from the CEPH collection (CEU).

RESULTS

科佰生物Cobioer

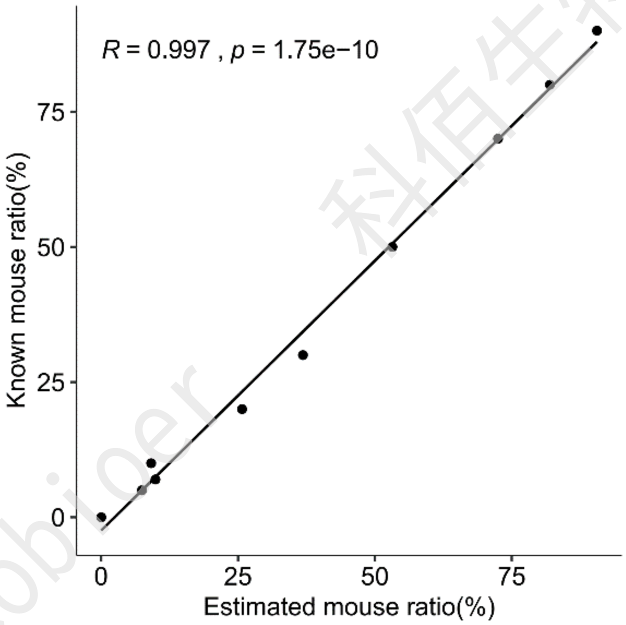


Comment: Mouse ratio is estimated based on 62 100-300bp human-mouse homologous segments that share identical flanking sequences and are amplified by same primers.

Conclusion: The mouse ratio in Sample PC-3-20260211 is 0%.

Technical note: This assay can estimate mouse ratio in a human-mouse mix with high accuracy (Figure 4A below, adopted from Reference [1]), and can reliably detect mouse ratio at ~0.1% [1].

科佰生物Cobioer



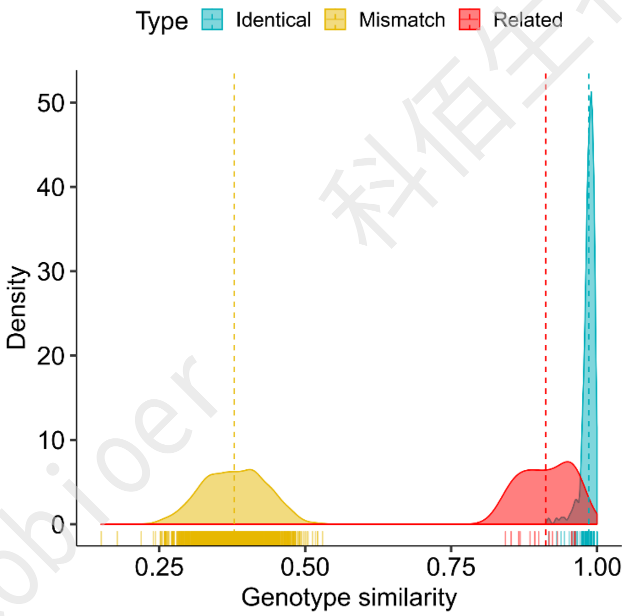
HUMAN SAMPLE AUTHENTICATION BY SNP METHOD

Reference	#Matched SNPs	#Nonmatched SNPs	#Total SNPs	Genotype similarity (%)
PC3	192	1	193	0.9948
HCC2108	108	87	195	0.5538
LN229	107	88	195	0.5487
NCIH1915	106	88	194	0.5464
DAOY	106	89	195	0.5436

Conclusion: PC-3-20260211 and PC3 are of same origin for the human part based on the high identity.

Technical note: Genotype similarities calculated from >200 SNPs are always >90% between identical samples even at the presence of minor contaminant. In contrast, genotype similarities between unrelated samples are almost always below 50%. If two samples are related, for example by deriving from same parental sample, their genotype similarity can range from about 80% to close to 100%, depending on their genetic divergence. Below: Figure 1A from Reference [1].

科佰生物Cobioer



HUMAN SAMPLE AUTHENTICATION BY STR METHOD

Conclusion: No matched human sample is detected.

Technical note: The values in parentheses under the column of “Test sample” are the proportion of reads supporting for this allele at the STR locus. When the STR similarity between the detected sample and the reference sample is greater than 80%, we consider these two samples to be a match. The STR information are from the Cellosaurus database.

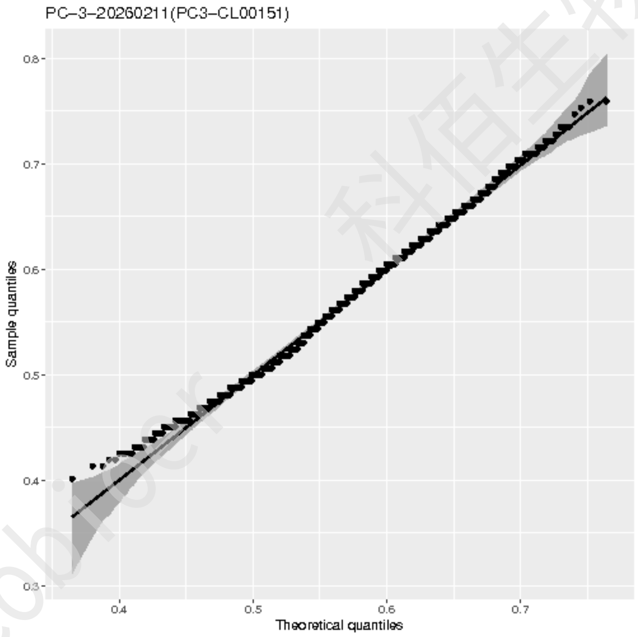
HETEROGENEITY RATIO

Conclusion: sample PC-3-20260211 has a low heterogeneity ratio with the above distribution of heterogeneity ratios for 227 SNPs.

Technical note: Heterogeneity ratio is a quantitative measure for genetic heterogeneity of a sample by using the >200 SNPs and is defined in Table 1 of Reference [1]. Heterogeneity ratio comes from genetic heterogeneity and contamination. High heterogeneity ratio, as well as a two/three-modal distribution, usually indicates possible contamination. See Reference [1] for details.

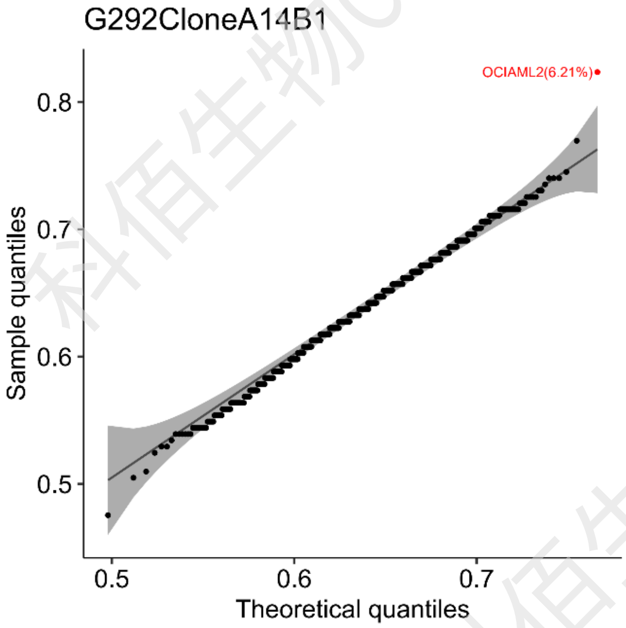
CONTAMINANT DETECTION FOR HUMAN SAMPLE

科佰生物Cobioer



Conclusion: No contaminant is detected for sample PC-3-20260211 .

Technical note: In this quantile-quantile plot, each dot is a reference cell line; theoretical and sample quantiles were calculated from a beta distribution fitted to genotype similarities between the test sample and all reference samples. The 99% confidence band is shaded. A reference sample was denoted as contaminant if (1) it had the highest genotype similarities, (2) its genotype similarity was above the 99% confidence upper bound in the quantile-quantile graph and (3) its P-value was <1.0E-6 in the fitted beta distribution. An example for contaminant is shown below (Figure 3D from Reference [1]).



SYNGENEIC MODEL IDENTIFICATION

Model	Characteristic SNPs	Characteristic SNPs detected	Detection ratio(%)
2666	0	0	NA
2PK3	0	0	NA
3T3J2	0	0	NA
3T3Swissalbino	0	0	NA

科佰生物Cobioer

Model	Characteristic SNPs	Characteristic SNPs detected	Detection ratio(%)
4T1	0	0	NA
A20	0	0	NA
ATDC5	0	0	NA
B16	0	0	NA
B16BL6	0	0	NA
B16F0	0	0	NA
B16F1	0	0	NA
B16F10	0	0	NA
BAF3	0	0	NA
BALB3T3	0	0	NA
BCL1clone5B1b	0	0	NA
C127I	0	0	NA
C1498	0	0	NA
C2C12	0	0	NA
CMT93	0	0	NA
CT26WT	0	0	NA
CloneM3	0	0	NA
Colon26	0	0	NA
D2N	0	0	NA
EG7Ova	0	0	NA
EL4	0	0	NA
EMT6	0	0	NA
EO771	0	0	NA
EOMA	0	0	NA
EPH41424	0	0	NA
FM3A	0	0	NA
FO	0	0	NA
GL261	0	0	NA
H22	0	0	NA
Hep551C	0	0	NA
Hepa16	0	0	NA
Hepa1c1c7	0	0	NA
J558	0	0	NA
J558L	0	0	NA
J774A1	0	0	NA
JC	0	0	NA
K7M2WT	0	0	NA

科佰生物Cobioer

Model	Characteristic SNPs	Characteristic SNPs detected	Detection ratio(%)
KLN205	0	0	NA
L1210	0	0	NA
L5178RLYR	0	0	NA
L5178SLYS	0	0	NA
L5178YTKclone372C	0	0	NA
LA4	0	0	NA
LLC	0	0	NA
MB49	0	0	NA
MBT2	0	0	NA
MC38-FDCC	0	0	NA
MC38-Kerafast	0	0	NA
MC38-Pharm1	0	0	NA
MC38-Pharm3	0	0	NA
MC38-NanjingGalaxy	0	0	NA
MC38-Pharm2	0	0	NA
MFC	0	0	NA
MLTC1	0	0	NA
MPC11	0	0	NA
MS1	0	0	NA
MTCM	0	0	NA
N18Hamprecht	0	0	NA
N1E115	0	0	NA
NFS60	0	0	NA
NIH3T3	0	0	NA
Neuro2a	0	0	NA
P388	0	0	NA
P388D1	0	0	NA
P3NS11Ag41NS1	0	0	NA
P3X63Ag8	0	0	NA
P3X63Ag8653	0	0	NA
P3X63Ag8U1	0	0	NA
P815	0	0	NA
Pan02	0	0	NA
RAG	0	0	NA
RAW2647	0	0	NA
RM1-ATCC	0	0	NA
RM1-SIBS	0	0	NA

科佰生物Cobioer

Model	Characteristic SNPs	Characteristic SNPs detected	Detection ratio(%)
Renca	0	0	NA
S491	0	0	NA
S49Thy1a	0	0	NA
SP20	0	0	NA
STO	0	0	NA
Sarcoma180	0	0	NA
TC1	0	0	NA
TRAMPC2WTE	0	0	NA
U14	0	0	NA
UGSM2	0	0	NA
WEHI164	0	0	NA
WEHI3	0	0	NA
WEHI3BD	0	0	NA
bEnd3	0	0	NA
mHypoEN6	0	0	NA

Comment: When >=80% of characteristic SNPs are detected, the matched mouse model will be assigned. The list of characteristic SNPs is available in Table S6 of Reference [1].

Conclusion: Sample PC-3-20260211 is not consisent with any mouse model in our database.

MICROORGANISM INFECTION DETECTION

Microorganism	Number of high-coverage bases
AAV	0
ADV	0
ADV-16	0
ADV-21	0
ADV-55	0
ADV-68	0
ADV-7	0
ADV-8	0
Adeno-associated_virus_9	0
B19V	0
BKV	0
Boone_Cardiovirus	0
Bovine_adenovirus_10	0
Bovine_adenovirus_2	0
Bovine_adenovirus_3	0

科佰生物Cobioer

Microorganism	Number of high-coverage bases
Bovine_adenovirus_4	0
Bovine_adenovirus_6	0
Bovine_adenovirus_7	0
Bovine_parvovirus_1	0
Bovine_parvovirus_2	0
Bovine_parvovirus_3	0
Bovine_respirovirus_3	0
Bovine_viral_diarrhea_virus_1	0
Bovine_viral_diarrhea_virus_2	0
Bovine_viral_diarrhea_virus_3	0
Burkholderia_gladioli	0
CMV	0
Corynebacterium_bovis	0
EBV	0
Ectromelia_virus	0
HAV	0
HBV	0
HCV	0
HFV1	0
HHV-5	0
HHV-6A	0
HHV-6B	0
HHV-7	0
HHV-8	0
HIV-1	0
HIV-2	0
HPV16	0
HPV18	0
HSV1	0
HSV2	0
HTLV-1	0
HTLV-2	0
HTLV-3	0
HTLV-4	0
Hantaan_orthohantavirus	0
Human_foamy_virus	0
JCV	0

科佰生物Cobioer

Microorganism	Number of high-coverage bases
Lactate_dehydrogenase-elevating_virus	0
Mammalian_orthoreovirus_1	0
Mammalian_orthoreovirus_2	0
Merkel_cell_polyomavirus	0
Minute_virus_of_mice	0
Mouse_parvovirus_1	0
Mouse_parvovirus_2	0
Mouse_parvovirus_3	0
Mouse_parvovirus_4	0
Mouse_parvovirus_5a	0
Murid_betaherpesvirus_1	0
Murid_betaherpesvirus_2	0
Murid_herpesvirus_3	0
Murine_adenovirus_1	0
Murine_adenovirus_2	0
Murine_chapparvovirus	0
Murine_coronavirus	0
Murine_hepatitis_virus	0
Murine_kobuvirus_1	0
Murine_norovirus	0
Murine_orthopneumovirus	0
Murine_poliovirus	0
Murine_polyomavirus	0
Murine_respirovirus	0
Murine_rotavirus	0
Murine_sapovirus	0
Porcine_circovirus_2	0
Porcine_circovirus_3	0
Porcine_parvovirus	0
Porcine_parvovirus_2	0
Porcine_parvovirus_3	0
Porcine_parvovirus_4	0
Porcine_parvovirus_5	0
Porcine_parvovirus_6	0
Porcine_parvovirus_7	0
Pseudomonas_aeruginosa	0
Rat_parvovirus_1	0

科佰生物Cobioer

Microorganism	Number of high-coverage bases
Rat_parvovirus_2	0
Rat_parvovirus_NTU1	0
Rotavirus_B	0
SV40	0
Saccharopolyspora_erythraea	0
Seoul_orthohantavirus	0
SinNombre_orthohantavirus	0
Theilers-like_virus_of_rats	0
Torque_teno_sus_virus	0
Torque_teno_sus_virus_1a	0
Torque_teno_sus_virus_1b	0
Torque_teno_sus_virus_k2a	0
Torque_teno_sus_virus_k2b	0
Treponema_pallidum	0
VSVG	0
Varicella-zoster_virus	0
Vesivirus	0
adeno-associated_virus_2	0

Comment: If the read depth is > 500, a base is accounted as a high-coverage one.

Conclusion: Sample PC-3-20260211 is negative for all detected microorganisms.

Technical note: Multiple pairs of primers are used to detect each microorganism. If the number of high-coverage bases for a virus exceeds 100, we determine that the sample is positive for this virus. See Reference [1] for details.

MYCOPLASMA CONTAMINATION DETECTION

Mycoplasma median coverage(universal region)	Human and mouse median coverage	Ratio of coverage
0	260	0

Conclusion: Sample PC-3-20260211 is mycoplasma negative.

Technical note: Six pairs of universal primers are used to detect mycoplasma contamination. The mycoplasma substrain is detected using specific primer of each substrain.

GENDER IDENTIFICATION FOR HUMAN SAMPLE

SNP-anchored Y chromosome segment	Read depth
hg19_chrY_14832620	0

科佰生物Cobioer

SNP-anchored Y chromosome segment	Read depth
hg19_chrY_15467824	0
hg19_chrY_15591537	0
hg19_chrY_2822023	0
hg19_chrY_7235632	0
hg19_chrY_21765821	0

Comment: If the sum of read depths of the 3 Y-chromosome SNPs >1000, the sample is identified as Male; If the sum if <300, the gender is predicted as Female; If the sum of depths is less than 1000 and more than 300, the gender is not assigned. Because Y chromosome is frequently lost in cancer genomes, it is advised to use microscope-based observation to confirm the gender prediction.

Conclusion: sample PC-3-20260211 is female.

GENETIC ADMIXTURE ANALYSIS

Comment: The human genetic admixture of sample PC-3-20260211 is not analysed.

REFERENCES

1. Chen,X., Qian,W., Song,Z., Li,Q. and Guo,S. (2020) Authentication, characterization and contamination detection of cell lines, xenografts and organoids by barcode deep NGS sequencing. NAR Genomics and Bioinformatics, 2, 3.
https://academic.oup.com/nargab/article-abstract/2/3/lqaa060/5893932.

科佰生物Cobioer