Figure S1

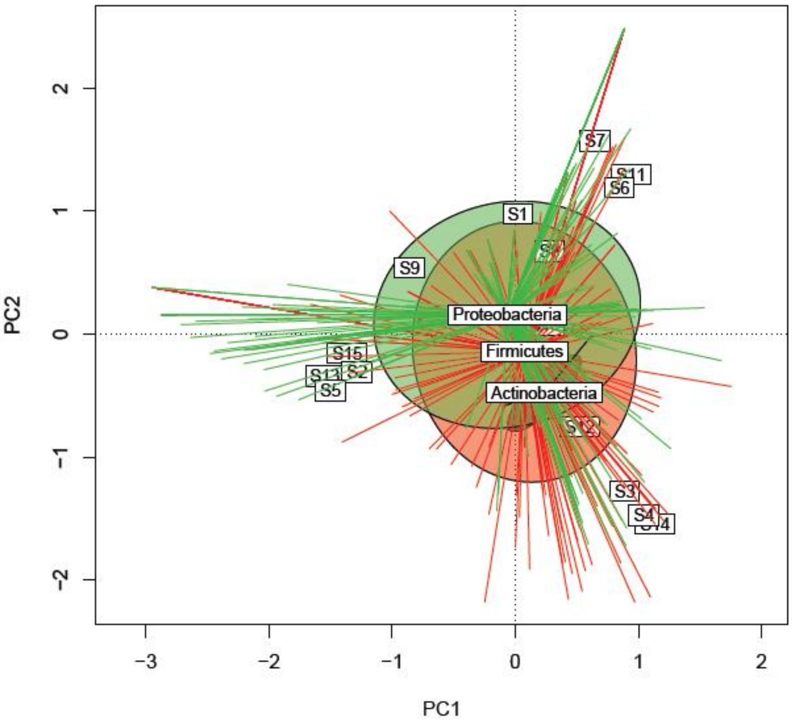

FIG S1. Principal component analysis results at the genus/species composition among samples from the human gut microbiota of athletes

Figure S2

```
LRH      ------------MTASPWDLRKVLDELKQDPQQYHETDVQVDPDAELAGVYRYIGAGGTV
LPL      ------------MAEQPWDLRRVLDEIKDDPKNYHETDVEVDPNAELSGVYRYIGAGGTV
PET      -----------MAKKDVYDLRKVLDELKKEPGQYHETNVEVDPDAELSGVYRYIGAGGTV
ERA      -------------MSTPYDLRKVLAELQEMPGEYHETNVEVDPHAELSGVYRYIGAGGTV
SGA      ------------MSEQPYDLRKVLEELKEIPGQYHETDVEIDPNAEISGVYRYIGAGGTV
ECL      ------------MQNPINDLRSAIALLQRHPGHYIETDHPVDPNAELAGVYRHIGAGGTV
ACH      MEQP---FPAFIDFPPINDLRSALARLQQVPGQLLHTDHPVDPHGELAGVYKRVGAGGTV
KIN      -----MNRTDHTANSPVIDLRSALARLEQSPDELISTDHEIDPRSELAGVYKRVGAGGTV
ATR      MEDNVMDRTDHAANSPVIDLRSALARLEAFPGELISTDQEVDPRSELAGVYKRVGAGGTV
AGL      ------------MNDTIHDLRSALDYLRDMPGQLLETDTEVDCDAEVSGVYRHVGAGGTV
ERY      -------------MKEPYDLRSALQQLEAEEGQLLITDKLTNTDGELAGVYRYIGGGGTL
COP      -----------MSKKKVRDLRSALELLQSIDGQLVETDVEVDPLGELSGVYRHVGAGGTV
BLA      ------------MAEKVRDLRSALALLEQMPDQLIETEVEVDPMAELAGVYRYVGAGGTV
FIR      ------------MADKVRDLRSALERLKTMEGQYIETDVEVDPMAELAGVYRYVGAGGTV
AHA      ------------MQDKVRDLRSALKRLKQMEGQYIETDVEVDPMAELAGVYRYVGAGGTV
LAC      ------------MQDKVRDLRSALKRLKQMEGQYIETDVEVDPMAELAGVYRYVGAGGTV
DLO      ------------MAAKVTDLRSALKMLEDMPGQLIETDVEVEPMAELSGVYRHVGAGGTV
IBA      -------------MNKVRDLRSALELLENMPGQLIETDVEVDPMAELSGVYRHVGAGGTV
CBU      ------------MSNKVYDLRSALELLKTLPGQLIETDVEVDPMAELAGVYRYVGAGGTV
HHA      ---------------MVEDLRSAVEELKKYENQIACTDTEVDSYAEVAGIYRYVGAGGTV
TSA      ---MKIKERVNKMANQIHDLRSAIEFLKQHENQIVYTNSEVDCEAEISGIYRYVGAGGTV
PAG      ---------MSGKESKVNDLRSALALLSQYDNELIYTDEPVDPVAELSGVYRYVGAHGTV
HAL      ----MANNADNKIKNPIHDLRSALEFLKNQPGELVSTNVEVDPCAELSGVYRYVGAGGTC
SBL      --------MKNKDKIPVHDLRSALELLKTLPGEYVETFTEVDPHAELSGVYRYVGAGGTC
KRA      -----MSNSENKNTSGVTDLRSAIELLKTLPGEYVETDTEVDPHAELSGVYRYVGAGGTC
KMI      -----MANSDNKTPSSVHDLRSALELLKTLPGEYVETDTEVDPHAELSGVYRYVGAGGTC
                           *** .:  :     .   *    :  .*::*:*: :*. **


LRH      ERPTQ-EGPAMMFNNVVGFPTTRVLIGLMASRKRVGKMFHQDYHTLGRFLNKAVLNPIQP
LPL      QRPTQ-EGPAMMFNNVKGFPDTRVLTGLMASRRRVGKMFHHDYQTLGQYLNEAVSNPVAP
PET      QRPTQ-EGPAMMFNNVKGFPDTRVLMGLMASRKRVGKMFHHDYHTLGQFLNDSVENPVDP
ERA      KRPTK-EGPAMMFNCVKGFPDTRVLIGIIASRDRVGKILHHDPKHLGRLLKDSVQNPVKP
SGA      ERPTQ-EGPAMTFNNIKGFPNVRVNIGTMASRKRVGHILHHDYKDLGHLLNKAVENPVKP
ECL      KRPTR-TGPAMMFNSVKGYPGSRILVGMHASRERAALLLGCVPSKLAQHVGQAVKNPVAP
ACH      KRPTR-LGPAMMFNHIKGYPDSRVLVGVMASRERVALLLDTTPDRLAERMGEAIEKAIDP
KIN      MRPTR-TGPAMMFENVKGYPGARVLVGLMAKRERVALLLDSRPEELGRRMGEAVLNGIAP
ATR      MRPTR-TGPAMMFENVKGYPGARVLVGLMAKRERVALLLDSKPEELGQRMGEAVLNGIAP
AGL      ARPTK-EGPAMVFNNVKGFDDAKVAIGLLSSRKRVAALLGMDEQYLGLQIGQKLAETIPP
ERY      QRPTQ-LGPAMLFTNIQNHPGSRVLIGLLGDRQRCASLLNTTSEQLPFLMNKAYSKRLDP
COP      MRPTQ-EGPAMIFNHVKGHPGARVAIGLLASRRRVGYLLDCEPQKLGFLLKDSVNNPIKP
BLA      KRPTK-EGPAMIFNRIKGHPDARVAIGLLASRKRVAALLDTQPENLGKMLCKSVENPIPP
FIR      KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKRVAALLDCKPEELGKKLYYSVDNPIPP
AHA      KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKHVAALLDCKPEELGKKLYHSVDNPIAP
LAC      KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKRVAALLDCKPEELGKKLYHSVDNPIAP
DLO      MRPTK-EGPAMIFNNVKGHPGARVAIGVLASRTRVGALLECDPKDLGKKLYHSVDNPIPP
IBA      QRPTK-EGPAMIFNNVKGHKDARVLIGLLASRKRVAALLGTEPEDLGKLLYRSVDNPIPP
CBU      QRPTK-EGPAMIFNNIKGHKDARVLIGLLASRRRVAALLDCEPENLGKLLYRSVDNPIAP
HHA      KRPTK-EGPALLFRNIKGFPDKQVLIGLLASRKRVGYLLDCPPDKLGFLLKDAAANPVMP
TSA      MRPTK-IGPAMLFNHVKDYEHSKVLIGLFASRERVGLMLGCEPDRLGFLLNDALNHPVDP
PAG      KRPTR-VGPAMVFNKIKGFNDMRVLIGLLSSRQRVARLFGTSPENLAFMLKDSVLNPIPP
HAL      QRPTRKNGPAMVFNKIKGFDDISVTIGLNGSRKRVGHFLNCAPERLGHLLKDSVKNAIQP
SBL      QRPTRKNGPAMVFHNIKGFRNTNVAIGLNGSRKRVGHFLNCAPEKLGFLLKDSVKHAIAP
KRA      QRPTRKNGPVMMFNKVKGFQDISVAIGLNGSRKRVSHFLNCAPEKLGHLLKDSVQNPIPP
KMI      QRPTRKNGPVMVFNKIKGFKDISVAIGLNGSRQRVSHFLQCEPAKLGHLLKDSVENAIAP
          ***:  **.: *  : ..    :  *  ..* : . ::      *   :     . : *
```

```
LRH        VTVE-ESAAPAHEVVAKASDPDFDIRKLVAAPTNTPQDAGPYITCGVVLGSNMAK-TMTD
LPL        ETVA-EADAPAHDVVYKATDEGFDIRKLVAAPTNTPQDAGPYITVGVVFGSSMDK-SKSD
PET        VMVD-EADAPTHEVVHKSTDADFDIRKLVAAPTNTPRDAGPYITVGVVYGSNPDK-TMSD
ERA        VKVA-KTDAPAQEVIHLATDEGFDIRKILAAPTNTEYDAGPYITMGVVFGSDPEK-TMSD
SGA        VKVS-KDQAPAQEVVHLATDDDFDIRKLIAAPTNTEYDAGPYITTGLVYGSTPDK-SMSD
ECL        VVVP-ASQAPCQEQVFYADDPDFDLRKLLPAPTNTPIDAGPFFCLGLVLASDPEDTSLTD
ACH        VVVP-REQAPCQEIVHRAEDPGFDLRTLLPAPTNTDEDAGPFFCLGLVLGSDPEE-GHAD
KIN        VVLK-GQSAPCQEQIFRADDPHFDLRNLLPAPTNTEEDAGPYFCLGLLLGSDPDN-GHTD
ATR        VVFK-EGSAPCQEQVFRADDPNFDLRHLLPAPTNTEEDAGPYFCLGLLLGSDPDN-GHTD
AGL        APIAEGKHIDCQEVVHLATDPDFDLRRLVPAPTNTPEDAGPYVTMGLIAGHSPFN-SDTD
ERY        VVVA---KGICQEIVHHREDEGFDIRRLLPAIKATKEDAGPYITMGLCYASEPET-KESD
COP        VVIP-HEQAKCQEVKYFADDPDFDIRKLVPAPTNTLEDAGPYITLGMCYASSPET-GESD
BLA        VDLE-G-DAPCQQVVHKAADPDFDLYKLVPAPTNTPDDAGPYITLGMCYATHPDT-GVHD
FIR        VEYQ-G-APPCQQVVHKAEDPDFNLYDLVPAPTNTPDDAGPYITLGMCYATHPDT-GVHD
AHA        VEYQ-G-DAPCQQVVHKVEDPDFNLYDLVPAPTNTPDDAGPYITLGMCYATHPDT-GVHD
LAC        VEYQ-G-DAPCQQVVHKVEDPDFNLYDLVPAPTNTPDDAGPYITLGMCYATHPDT-GVHD
DLO        VLTE-EA-APCQEVVHKATDPDFDLYKLVPAPTNTPVDAGPYITLGMCYASHPDT-GVSD
IBA        VLTD--KTPLCQEVVHRATDEDFDLYKLVPAPTNTPDDAGPYITLGMCYASHPDT-GCSD
CBU        VLTD-AKLPLCQQVVHKATDPDFDLNKLVPAPTNTPDDAGPYITLGMCYASHPDT-KFSD
HHA        EFTE-SGDVPCQEEIHYASDEGFDIRKILPAPQNTEEDAGPYITMGMCYASDPVT-GDGD
TSA        IEIP-QAKAKCQEVVHLSTDEGFDIRKILPAPKNTLEDAGPYITMGLCYASDPDT-KAGD
PAG        IVIP-REHAVCQEVVHLATDPDFDILKILPTPTNTPEDAGPYFTLGMCYAADPET-GEHD
HAL        VTVS-ADKAVCQQVTHFAEDADFDLRKLLPAPTNTEEDAGPYITMGLCYASDPDT-HESD
SBL        VDVS-TGNAVCQEVVHLATDPDFDLRKLLPAPTNTEEDAGPYITMGLCYASDPQT-HESD
KRA        IL-T-KDNAVCQQVVHLASDTNFDLRKLLPAPTNTEEDAGPYITMGLCYASDPET-HESD
KMI        VM-A-QGAAVCQQVVHLASDADFDLRKLLPAPTNTEEDAGPYITMGLCYASDPET-HESD
                      ::      *  *::  :: :   *  ****:.  *:  .         *


LRH        VTIHRMVLEDKDTLGIYIMPGGRHIGHFAEEYEKANKPMPVTINIGLDPAITIGATFEPP
LPL        VTIHRMVLEDKDKLGIYIMPGGRHIGAFAEEYEKANKPMPITINIGLDPAITIGATFEPP
PET        VTIHRMVLEDKDKLGIYIMPGGRHIGKFAEEYEKMNKPMPITINIGLDPAITIGATFEPP
ERA        VTIHRMVLEDEDTIGMYIMPGGRHIGHFQKQYEALNKPMPITINIGLDPAISIGTTFEPP
SGA        VTIHRMVLEDKDTIGIYIMPGGRHIGAFLSEYQKLNKPMPITINIGLDPAILIGATFEPP
ECL        VTIHRLCVQERDELSMFLAAG-RHIEVFRKKAEAAGKPLPVTINMGLDPAIYIGACFEAP
ACH        VTIHRLCVQGRDRLSIFFAPD-RHIDKFRQKAEAAGKPLPVTINMGLDPAILIGSCFEAP
KIN        VTIHRLCVQGRDELSVFFAPG-RHIDAFRAKAEERGEALPITINMGLDPAIPIGACFEAP
ATR        VTIHRLCVQGRDELSVFFAPG-RHIDAFRAKAEERGEALPITINMGLDPAIPIGACFEAP
AGL        VTIHRLCIESKDTMGMWITPGSRHLGAFFEQWKEKGEDMPVTVSIGLDPAVYMCAGFEAP
ERY        VTIHRLCLQSKDEMTMFFTPGIRHLDVFRKKAEAKNKALPISISIGVDPAIEIAACFEPP
COP        VTIHRMCFQSKDEISIFLQPGARHIGYFRELAEAKGEALPISISIGVDPAIEIASCFEAP
BLA        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQRLPISISIGVDPAIEIGSCFEPP
FIR        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPAIEIGSCFEAP
AHA        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPAIEIGSCFEAP
LAC        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPAIEIGSCFEAP
DLO        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQRLPISISIGVDPAIEIGSCFEPP
IBA        VTIHRLCIQSKDELSIFFTPGARHIGAMAERAEQLGQKLPISISIGVDPAIEIGSCFEPP
CBU        VTIHRLCIQGKDELSIFFTPGARHIGAMAERAEELGQNLPISISIGVDPAIEIGSCFEPP
HHA        VTIHRLCLQSADEMTMFFTPGVRHLDAFREKAEREGVNLPISISIGVDPAIEIASCFEPP
TSA        ITIHRLCLQSKDELTMFFTPGVRHIDAFRKKAEGNQNNLPISISIGVDPAIYVTSCFEPP
PAG        VTIHRLCVQSKDEISMYFVPG-RHLDTFRQKAEAAGKPLPITISIGVDPAIEIGACFEPP
HAL        ITIHRLCVQSRDELSMWLTPG-RHIDAFRMKAEAQGKPLPISISIGVDPAIEIAACFEPP
SBL        ITIHRLCVQSKDELTMWLTPG-RHIDAFRIKAEAAGKALPISISIGVDPAIEVAACFEPP
KRA        ITIHRLCVQSRDELSMWLTPG-RHIDAFRMKAEAAGKPLPISISIGVDPAIEIAACFEPP
KMI        ITIHRLCVQSRDELSMWLTPG-RHIDAFRMKAEAAGQPLPISISIGVDPAIEIAACFEPP
           :****: .:  * : :::  . **:  :    :     :*:::.:*:***: : : ** *
```

```
LRH      TTPLGYDELGVAGAIRQEPVQLVQAVTVDEKAIARSEFTLEGYIMPNTRIQEDINTHTGK
LPL      TTPFGYNELGVAGAIRNQAVQLVDGVTVDEKAIARSEYTLEGYIMPNERIQEDINTHTGK
PET      TTPLGYNELGVAGAIRQEPVQLVKGLTVDEKAIARSEYTLEGYIMPNERIQEDINTHTGK
ERA      TTPLGYNELWVAGALRQEPVQLVDGVAVDEVGIARSEFIIEAEILPHETIQEDINTNTGK
SGA      TTPLGYNELWVAGALRNEPVQLVDSIAVDEVGIARSEFIIEGEILPNETIQEDINTHTGH
ECL      TTPFGYNELGVAGALRQQPVELVQGVAVKEKAIARAEIIIEGELLPGVRVREDQHTNTGH
ACH      TTPLGYDELKIAGGLRGRPVELVEAVSIGQKAIARAEVVIEGEILPHERLREDINTDSGR
KIN      TTPLGFDELTVAGGLRGRAVELVDAVTVKERSIARAEIVIEGEILPGRRIREDVNTNTGH
ATR      TTPLGFDELTVAGGLRGRAVELVDAVTVKERSIARAEIVIEGEILPGRRIQEDVNTNTGH
AGL      TTPLGYNELQIAGGIRGRAVELAPALTVPETCIAHAEYVIEGYLSITDTVREDVNTNTGK
ERY      TTPLGFNELSIAGGIRNKPVELCKCVSIDEYAIAHSEYVIEGELLPGKRMREDQHTGLGK
COP      TTPLGYDELQVAGAIRNEPVELVECLTINERAIANAEYVIEGEILPGKRIREDIHSNTGK
BLA      TTPLGYDELSVAGALRKEPVELCKCLTVNERAIANAEYVIEGEVIPGVRVKEDQNSNTGY
FIR      TTPLGYDELAVAGALRNEPVELCKCLTVNEMAIANAEYVIEGEVILNVRVQEDQNSHTGY
AHA      TTPLGYDELSVAGALRNEPVELCKCLTVNEMAIANAEYVIEGEVIPNVRVQEDQNSHTGY
LAC      TTPLGYDELSVAGALRNEPVELCKCLTVNEMAIANAEYVIEGEVIPNVRVQEDQNSHTGY
DLO      TTPMGYDELAVAGALRGEPVRLCKCLTVNERAIANAEYVIEGEVVPNVRVQEDQNSHTGY
IBA      TTPMGYDELAVAGALRGEAVELCNCLTVNEKAIANAEYVIEGEVIPNVRVQEDQNSNTGY
CBU      TTPLGYDELSVAGALRGKPVELCKCITVNERAIANAEYVIEGEVIPNLRVQEDKNSNTGY
HHA      TTPLGFNELSIAGALRKKPVRMAKCRTIDEYAIANAEFVIEGELVAGKRMREDINTNTGK
TSA      ATPLGFNELSIAGGLRGKPVELTKCLTIDEMAIANAEYVIEGELVFGQRMREDINTNTGK
PAG      TTPLGFDELSVAGSLRNQAVELVDCLTVNARGIANAEIVIEGELVPNYRVREDQNTNTGK
HAL      TTPLGFDELSIAGALRGHAVEMVQCKTINEKAIAHAEIVIEGELLPNLRVREDQNTNTGK
SBL      TTPLGFNELSIAGALRGRAVEMVQCKTINEKAIAHAEIVIEGELLPDVRMQEDINTHTGR
KRA      TTPLGYDELSVAGALRGKAVEMVQCKTINERAIAHAEIVIEGELLPNVRLREDQNTNTGK
KMI      TTPLGFNELSIAGALRGKAVEMTQCKTINEKAIAHAEIVIEGELRPNARVREDQNTHTGR
         :**:*::** :**.:* . *.:    ::    **.:*  :*. :     ::** ::  *


LRH      AMPEFPGYDGDANPALQVIKVTAVTHRRDHPIMQSVIGPSEEHVSMAGIPTEASILQLVD
LPL      AMPEFPGYDGDANPALQVIKVTAVTHRK-NAIMQSVIGPSEEHVSMAGIPTEASILQLVN
PET      AMPEFPGYDGDANPALQVIKVTAVTHRK-DPIMQSVIGPSEEHVSMAGIPTEASILSLTN
ERA      AMPEFPGYNGDANPAVNVVKVKAITHRKDRPIMQTTIGPSEEHVSMAGIPTEASILDLVD
SGA      AMPEFPGYNGPANPALNVIKVKAVTHRKDNPIMQTTIGPSEEHVSMAGIPTEASILNLVD
ECL      AMPEFPGYCGEANPSLPVIKVKAVTMRN-HAILQTLVGPGEEHTTLAGLPTEASIRNAVE
ACH      AMPEFPGYTGPANPSLPVIRVTAVTMRR-QAILQTLVGPGEEHTNLAGIPTEASIRNAVE
KIN      AMPEFPGYNGPANPALPVIKVKAVTMRK-NAILQTLVGPGEEHVNLAGIPTEASIFNACD
ATR      AMPEFPGYNGPANPSLPVIKVTAVTMRK-NAILQTLVGPGEEHVNLAGIPTEASIFNACD
AGL      AMPEFPGYTGEAKPELPVIHVTAVTHRE-HPIMQSCIGPSHEHVSMAGIPTEASIYKMVE
ERY      AMPEFPGYCGEANPSLPVIRVKAITHRK-NPIMRVCLGPGEEHVNLAGIPTEASILSLCE
COP      AMPEFPGYCGPANPELPIIKVKAVTTRK-NPIMQTCIGPSEEHVSMAGIPTEASILTMVD
BLA      AMPEFPGYTGPASDQCWLIKVTAVTHRE-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
FIR      AMPEFPGYTGPASSQCWLIKVKAVTHRE-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
AHA      AMPEFPGYTGPASSQCWLIKVKAVTHRE-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
LAC      AMPEFPGYTGPASSQCWLIKVKAVTHRE-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
DLO      AMPEFPGYTGPASDQCWMIKVTAVTHRK-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
IBA      AMPEFPGYTGPASDQCWMIKVTAVTHRK-NPIMQTCIGPSEEHVSMAGIPTEASIYGMVE
CBU      AMPEFPGYTGPASDQCWMIKVKAVTHRE-NPIMQTCIGPSEEHVSMAGIPTEASIYGMIE
HHA      AMPEFPGYTGHVSGELPVIKIKAVTHRK-NPMIQTCIGPSEEHVSMAGIPTEASILTMEE
TSA      AMPEFPGYTGVADQALPVIKVKAITHRK-NPIMQTCIGPSEEHVNLAGIPTEASILNMVE
PAG      AMPEFPGYTGAAQAEVPVIKVKAITHRR-HPILQTCIGPSDEHTNLAGIPTEASILQMIE
HAL      AMPEFPGYTGAAKDALPVIKVKAVTHRY-HPIWRTTIGPGEEHVNMAGIPTEASILDMVE
SBL      AMPEFPGYTGEAKAAIPVIKVKAVTHRV-NPIWRTTLGPGEEHVNMAGIPTEASILDMVE
KRA      AMPEFPGYTGEAKDALPVIKVKAVTHRY-NPIWRTTVGPGEEHVNMAGIPTEASILDMVG
KMI      AMPEFPGYTGEAKEAIPVIKVKAVTHRI-NPIWRTTVGPGEEHVNMAGIPTEASILDMVE
         ******** * ..    ::::.*:* *    : :  :**..**..:**:******
```

```
LRH         RAIPGKVKNVYNPPAGGGKLMTIMQIHKDNPADEGIQRQAALLAFSAFKELKTVWLVDDD
LPL         RAIPGKVTNVYNPPAGGGKLMTIMQIHKDNEADEGIQRQAALLAFSAFKELKTVILVDED
PET         RAIPGKVLNVYNPPAGGGKLMTIMQIHKDNAADEGIQRQAALLAFSSFKELKTVFLVDED
ERA         KAIPGKVTNVYNPPAGGGKLMSILQIHKESEADEGIQRQAAILALSAFKELKTVILVDDD
SGA         KAIPGKVLNVYNPPAGGGKLMTIMQIRKENPADEGIQRQAALLAFSSFKELKTVILVDED
ECL         EAIPGFLQNVYAHTAGGGKFLGILQVKKRQPSDEGRQGQAALIALATYSELKNIILVDED
ACH         RAMPGLLKNVYAHTAGGGKLLAVLQVAKRRPGDEGRQRQAALIALAVYRELKNVILVDED
KIN         KALPGFVKNVYAHSAGGGKLLAILQVCQRSAGDVGKARQAALIALAVYRELKNVIIVDDD
ATR         KALPGFVKNVYAHSAGGGKLLAILQVCQRSAGDAGKARQAALIALAVYRELKNIIIVDDD
AGL         TAIPGRLLNVHAAPCGGGKFVAVLQFKKSSINDEGRQRNAAMLAFSAFSELKHVFLVDED
ERY         KAMPGNVRNVHCASSGGGKYIAVLQFVKRMESDEGRQRQAALLAFSAFSELKHVFLVDED
COP         KAMPGKLQNVYCASSGGGKYIAVMQFKKSVPSDEGRQRQAALLAFSAFAELKHVFLVDED
BLA         KAMPGRLQNVYCSSAGGGKYMAVLQFKKLSQSDEGRQRQAALLAFSAFSELKNVFLVDED
FIR         KAMPGRLQNVYCGSAGGGKYMAVLQFKKREASDEGRQRQAALLAFSAFSELKHVFLVDED
AHA         KAMPGRLQNVYCGSAGGGKYMAVLQFKKREASDEGRQRQAALLAFSAFSELKHVFLVDED
LAC         KAMPGRLQNVYCGSAGGGKYMAVLQFKKREASDEGRQRQAALLAFSAFSELKHVFLVDED
DLO         KAMPGRLQNVYCASPGGGKYMAVLQFKKLTASDEGRQRQAALLAFSAFSELKHVFLVDED
IBA         KAMPGRLQNVYCSSAGGGKYMAVLQFKKSVASDEGRQRQAALLAFSAFSELKNIFIVDED
CBU         KAMPGRLQNVYCCSSGGGKFMAVLQFKKTVASDEGRQRQAALLAFSAFSELKNIFLVDED
HHA         KALAGRVKNVYAHRSGGGKYMAVIQFAKQVPSDEGRQRQAALIALTAFPELKHVIVVDED
TSA         RAMPGRVLNVYAHSSGGGKYLAILQFKKSIPSDEGRQRQAALVAFSAFPELKHVILVDED
PAG         RALPGFVQNVHCPSPGTGKYLAVLQVKKRFAVDEGRQRQAALLAFSAFSELKHVMLVDED
HAL         RAMPGKLLNVFAHSAGGGKLLAVMQFKKSAPVDEGRQRQAALLAFSAFSELKHVILVDED
SBL         RAMPGKLLNVFAHSAGGGKLLAVMQFKKSSVNDEGRQRQAALLAFSAFPELKHVILVDED
KRA         RAMPGKLLNVFAHSAGGGKLLAVMQFKKFSPADEGRQRQAALLAFSAFPELKHVILVDED
KMI         RAMPGKLLNVFAHSAGGGKLLAVLQFKKSSPADEGRQRQAALLAFSAFPELKHVILVDED
             *: * : **.    * ** : ::*. :    * *   :**::*:: : *** : :**:*
```

```
LRH         VDIFDMNDVVWTMNTRFQGDQDIMVLPGMRNHPLDPSERPQYDPKSIRVRGMSSKTVIDG
LPL         VDIFDMNDVIWTMNTRFQADQDLMVLSGMRNHPLDPSERPQYDPKSIRFRGMSSKLVIDG
PET         VDIFDMNDVVWTINTRFQADQDIMVLPGMRNHPLDPSERPEYDPKSIRTRGMSSKLVIDG
ERA         VDIFDMNDVMWTLNTRFQGDQDILVLPGMRNHPLDPSERPEYDPKSIRFRGMSSKTILDG
SGA         VDIFDMNDVMWTINTRFQAHKDIMSLEGMRNHPLDPSERPEYSPEHIRVRGMSSKLVLDG
ECL         VDIFDSDDILWAMTTRMQGDVSITTLPGIRGHQLDPSQSPDYST-SIRGNGISCKTIFDC
ACH         VDPFDSDDVLWAMQTRYQGDVDTIFVPGVPGHVLDPSQVPEYSP-SIAARGLTCKTIFDC
KIN         VDLFDSNDVLWAMQTRYQGNVDTMFLPGVTGHVLDPSQVPEYDP-SIPAKGVSCKTIFDC
ATR         VDLFDSNDVLWAMQTRYQGNVDTMFLPGVTGHVLDPSQVPDYDP-SIPAKGVSCKTIFDC
AGL         VDIFDMSDVMWAMTTRFQADTGLITIPGAHCHVLDPSNDPAFAP-SIRVHGIACKAIFDC
ERY         VDIYDMKDVLWAMTTRFQSDIDCISIPGVRCHPLDPSNDTTYDP-SIRDRGIACKTIFDC
COP         VDPFDMKDVMWAMTTRFQADLDLITIPGVQCHPLDPSNQPEYSA-HIRARGVACKAIFDC
BLA         VDCFDMNDVLWAMNTRFQGDVDVITIPGVRCHPLDPSNDPAFSP-SIRDHGIACKTIFDC
FIR         VDCFDMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSP-SIRDHGIACKTIFDC
AHA         VDCFDMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSS-SIRDHGIACKTIFDC
LAC         VDCFDMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSS-SIRDHGIACKTIFDC
DLO         VDCFDMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPSCSG-SIRDHGIACKTIFDC
IBA         VDCFDMNDVLWAMNTRFQGDVDVITIPGVRCHPLDPSNDPSFSP-SIRDHGIACKTIFDC
CBU         VDCFDMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPDYSP-SIKNHGIACKTIFDC
HHA         VDIFDSDDVLWALNTRYQGDVDTITIPGVRCHPLDPTEGPEYNP-MLKDRGISCKTIFDC
TSA         VDIFDSNDVLWALTTRYQGDLDTVFIPGVRCHPLDPSSSNLYHP-SILSTGIACKTIYDC
PAG         VDVFDLSDVMWAMTTRYQGDVSTVFIPGVRCHPLDPSSDPAFSP-SVRDHGITCKTIFDC
HAL         VDIFDSDDVMWAMQTRYQGDVDTVFIPGVRCHPLDPSQMPDYSP-SILQEGMSCKTIFDC
SBL         VDIFDTDDVLWAMQTRYQGDIDTITIPGVRCHPLDPSQVPEYSP-FITQQGMTCKTIFDC
KRA         VDIFDSDDVLWAMQTRYQGDVDTVTIPGVRCHPLDPSQIPAYSP-SILQQGMSCKTLFDC
KMI         VDIFDSDDVLWAMQTRYQGDVDTIVIPGVRCHPLDPSQVPEYSP-SILQQGMSCKTIFDC
            ** :* .*::*:: ** *.. .   : *   * ***:.        :   *::.* : *
```

```
LRH    TVPFDMRDQFKRAAFKKVS--DWQKYLK-----
LPL    TVPFDMKDQFERAQFMKVA--DWEKYLK-----
PET    TVPFDMKDQFERAQFKEVK--DWEKYLK-----
ERA    TVPFDLKDDFIRAEFKEVP--DWEKYLK-----
SGA    TVPFDMKDQFERAKFKEVP--DWKKYLD-----
ECL    TVPWALKARFERAPFMEVDPTPWAPELFSDKK-
ACH    TAPWHLRERFERAKFRDVDPHPFAPELFPQARS
KIN    TYPWKLKEHFVRAQFRDVDPHPFAPSIFPKAGM
ATR    TYPWKLKEHFVRAQFREVDPHPFAPSIFPKAGM
AGL    TVPYDLKDEFQRCRFLDIDQDKWAAELAH----
ERY    TVPFSQKERFQRAAFQEVEASQWLK--------
COP    TVPFDQKARFERAKFMDVDPKCWLPDMLK----
BLA    TVPYDLKEDFHRARFMELDPEKWLKK-------
FIR    TVPYDQKERFKRARFMEVDPEHWL---------
AHA    TVPYDQKERFKRARFMEVDPEHWLS--------
LAC    TVPYDQKERFKRARFMEVDPEHWLS--------
DLO    TVPYDQKDRFKRAQFMDVDPEHWVK--------
IBA    TVPFEQKERFKRARFMDVDPEHWLKDIK-----
CBU    TVPFHMKERFKRAKFMEVDPEHWL---------
HHA    TVPYGLKDRFQRSKFKEVNMEDYEIRPLV----
TSA    TVPFIQKERFKRSSFKEVNLSDYDIKPLK----
PAG    TVPYNLKANFHRSEFLEVDVNRFIPGFNQK---
HAL    TVPFHLKAHFERSRFKEADVKRFLPDFE-----
SBL    TVPFHLKTHFERSTFKEVDVKRFLPDFE-----
KRA    TVPFHLKSHFERSKFKEVDVKRFLPDFE-----
KMI    TVPFHLKHTFQRSRFKEVDVKRFLPDFE-----
       * *:  :  * *. * .     :
```

FIG S2. Comparison of amino acid sequences of subunit C of putative gallate decarboxylase proteins from bacteria identified by HTS from athlete faecal microbiota. Multiple alignments were done using the programs ClustalOmega after retrieval of sequences from BLAST homology searches. The C subunits are from *Acetobacter tropicalis* (ATR) (WP_006559787.1), *Actinomyces glycerinitolerans* (AGL) (WP_073329259.1), *Anaerostipes hadrus* (AHA) (WP_009204323.1), *Azotobacter chroococcum* (ACH) (WP_052264016.1), *Blautia* sp. KLE 1732 (BLA) (WP_021650680.1), *Clostridium butyricum* (CBU) (WP_058371993.1), *Coprobacillus* (COP) (WP_008787659.1), *Dorea longicatena CAG:42* (DLO) (CDE20644.1), *Enterobacter cloacae* subsp. *cloacae* ATCC 13047 (ECL) (YP_003612445.1), *Enterococcus* (ERA) (WP_010743655.1), *Erysipelotrichaceae* bacterium 3_1_53 (ERY) (EFP59895.1), Firmicutes bacterium CAG:270 (FIR) (CDD72954), *Hafnia* *alvei* ATCC 51873 (HAL) (WP_004092226.1), *Hungatella hathewayi* (HHA) (WP_006771947.1), *Intestinibacter bartlettii* (IBA) (WP_007285641.1), *Klebsiella michiganensis* KCTC 1686 (KMI) (AEX02211.1), *Komagataeibacter intermedius* (KIN) (WP_039733191.1), *Kosakonia* (KRA) (WP_071921386.1), *Lachnospiraceae* bacterium 5_163FAA (LAC) (EFV16561), *Lactobacillus rhamnosus* (LRH) (WP_005712295.1), *Lactobacillus plantarum* ATCC 14917T (LPL) (D7VDD5), *Pantoea agglomerans* strain FDAARGOS_160 (PAG) (AMG60167.1), *Pediococcus ethanolidurans* (PET) (WP_057806460.1), *Shimwellia blattae* (SBL) (WP_002441866.1), *Streptococcus* (SGA) (WP_003065832.1), and *Turicibacter sanguinis* (TSA) (WP_040763984.1). Residues that are identical (*), conserved (:) or semiconserved (.) in all sequences are indicated. Dashes indicated gaps introduced to maximize similarities. Degenerate primers were designed on the conserved domains highlighted.