Supplementary Data 2

Fig. 1S Gene cluster annotation for the erythromycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AY661566) with *ab initio* gene predictions using the SEARCHPKS and ***ClustScan*** annotation programs

Fig. 2S Gene cluster annotation for the niddamycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AF016585) with *ab initio* gene predictions using the SEARCHPKS and ***ClustScan*** n annotation programs

Fig. 3S Gene cluster annotation for the tylactone producing polyketide synthase, comparing the original sequence annotation (Genbank accession number SFU78289) with *ab initio* gene predictions using the SEARCHPKS and ***ClustScan*** annotation programs

Fig. 4S Gene cluster annotation for the rifamycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AF040570) with *ab initio* gene predictions using the SEARCHPKS and ***ClustScan*** annotation programs

Supplimentary Data 2 Figure 1S: Gene cluster annotation for the erythromycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AY661566) with *ab initio* gene predictions using the SearchPKS and ClustScan annotation programs.

	EryAI	EryAII	EryAIII
Erythromycin annotated	AT ACP KS AT KR ACP KS AT KR ACP C3p C3 S|R C3 R|S	KS AT kr ACP KS AT DH ER KR ACP C3 S|x C3 x|x	KS AT KR ACP KS AT KR ACP TE C3 R|S C3 R|S
SearchPKS	AT ACP KS AT KR ACP KS AT KR ACP ACP CX C3 C3	KS AT KR ACP KS AT DH ER KR ACP C3 C3	KS ACP AT KR ACP KS AT KR ACP TE CX C3
Clustscan	AT ACP KS AT KR ACP KS AT KR ACP C3p C3 S|R C3 R|S	KS AT kr ACP KS AT DH ER KR ACP C3 S|x C3 x|R	KS AT KR ACP KS AT KR ACP TE C3 R|S C3 R|S

Legend:

All domain names written in capital letters are active.
Domain names written in small letters are inactive.
Domain names in red differ from annotated clusters.

AT:
C3p - propionyl
C3 - methylmalonyl
CX - not predictable

KR stereochemistry:
Me|OH
R - Rectus
S - Sinister
x - not predictable

POTENTIAL POLYKETIDE DOMAIN ORGANISATION OF YOUR SEQUENCE

Click any image below

Your sequence

APPROX PRODUCT CHEMISTRY :

--Loading-(CH0-CHR)-(CH0-CHR)-ACP alone-
R=xxx | R=CH3 | R=CH3 | ----- |

Your sequence

APPROX PRODUCT CHEMISTRY :

(CH0-CHR)-(CH2-CHR)-
R=CH3 | R=CH3 |

Your sequence

APPROX PRODUCT CHEMISTRY :

(xxxxxxx)-(xxxxxxx)-(CH0-CHR)-PKSend-TE-
| R=xxx | R=CH3 |----- |

Supplimentary Data 2 Figure 2S: Gene cluster annotation for the niddamycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AF016585) with *ab initio* gene predictions using the SearchPKS and ClustScan annotation programs.

	nidA1													nidA2					nidA3										nidA4				nidA5				
Niddamycin																																				
annotated	KSQ	AT	ACP		KS	AT	KR	ACP		KS	AT	DH	KR	ACP	KS	AT	DH	KR	ACP	KS	AT	kr	ACP		KS	AT	DH	ER	KR	ACP	KS	AT	KR	ACP	KS	AT	KR	ACP	TE
		C2a				C2	x|R				C2		x|x			C2		x|x		C3	x|x				EM			x|x			MM	R|S			C2	x|S		
SearchPKS	KS	AT	ACP		KS	AT	KR	ACP		KS	AT	DH	ACP	KR	ACP	KS	AT	DH	KR	ACP	KS	AT	KR	ACP		KS	AT	DH	ER	KR	ACP	KS	AT	KR	ACP	KS	AT	KR	ACP	TE
		C2				C2					C2					C2					C3					CX						EM			C2				
ClustScan	KSQ	AT	ACP		KS	AT	KR	ACP		KS	AT	DH	KR	ACP	KS	AT	DH	KR	ACP	KS	AT	kr	ACP		KS	AT	DH	ER	KR	ACP	KS	AT	KR	ACP	KS	AT	KR	ACP	TE
		C2a				C2	R|R				C2		R|R			C2		R|R		C3	S|x				EM			R|R			EM	R|S			C2	R|S		

Legend:

All domain names written in capital letters are active.
Domain names written in small letters are inactive.
Domain names in red differ from annotated clusters.

AT:
C2 - malonyl
C2a - acetyl
C3 - methylmalonyl
MM - methoxymalonyl
EM - ethylmalonyl
CX - not predictable

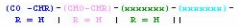
KR stereochemistry:
Me|OH
R - Rectus
S - Sinister
x - not predictable

POTENTIAL POLYKETIDE DOMAIN ORGANISATION OF YOUR SEQUENCE

Click any image below

Your sequence

APPROX PRODUCT CHEMISTRY :

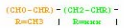


Your sequence

APPROX PRODUCT CHEMISTRY :

Your sequence

APPROX PRODUCT CHEMISTRY :

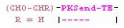


Your sequence

APPROX PRODUCT CHEMISTRY :

Your sequence

APPROX PRODUCT CHEMISTRY :



Supplimentary Data 2 Figure 3S: Gene cluster annotation for the tylactone producing polyketide synthase, comparing the original sequence annotation (Genbank accession number SFU78289) with *ab initio* gene predictions using the SearchPKS and ClustScan annotation programs.

Tylactone	tylG1	tylG2	tylG3	tylG4	tylG5
annotated	KSQ AT ACP KS AT KR ACP KS AT DH KR ACP C3p C3 R|R C3 x|x	KS AT DH KR ACP C2 x|x	KS AT kr ACP KS AT DH ER KR ACP C3 x|x EM x|x	KS AT KR ACP C3 R|S	KS AT KR ACP TE C2 x|S
SearchPKS	KS AT ACP KS AT KR ACP KS AT DH KR ACP C3 C3 C3	KS AT DH KR ACP C2	KS AT KR ACP KS AT DH ER KR ACP C3 EM	KS AT KR ACP C3	KS AT KR ACP TE C2
ClustScan	KSQ AT ACP KS AT KR ACP KS AT DH KR ACP C3p C3 R|R C3 R|R	KS AT DH KR ACP C2 R|R	KS AT kr ACP KS AT DH ER KR ACP C3 S|x EM R|R	KS AT KR ACP C3 R|S	KS AT KR ACP TE C2 R|S

Legend:

All domain names written in capital letters are active.
Domain names written in small letters are inactive.
Domain names in red differ from annotated clusters.

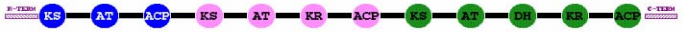
AT:
C2 - malonyl
C3 - methylmalonyl
C3p - propionyl
EM - ethylmalonyl
CX - not predictable

KR stereochemistry:
Me|OH
R - Rectus
S - Sinister
x - not predictable

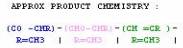
POTENTIAL POLYKETIDE DOMAIN ORGANISATION OF YOUR SEQUENCE

Click any image below

Your sequence



APPROX PRODUCT CHEMISTRY :

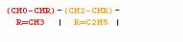


Your sequence

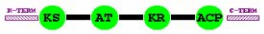
APPROX PRODUCT CHEMISTRY :

Your sequence

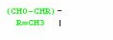
APPROX PRODUCT CHEMISTRY :



Your sequence



APPROX PRODUCT CHEMISTRY :



Your sequence

APPROX PRODUCT CHEMISTRY :

Supplimentary Data 2 Figure 4S: Gene cluster annotation for the rifamycin producing polyketide synthase, comparing the original sequence annotation (Genbank accession number AF040570) with *ab initio* gene predictions using the SearchPKS and ClustScan annotation programs.

Rifamycin – rifA

	Module	Module	Module	Module
annotated	ACP	KS AT dh KR ACP	KS AT ACP	KS AT kr ACP
		C3 x|R	C2	C3 x|x
SearchPKS	ACP	KS AT DH KR ACP	KS AT ACP	KS AT KR ACP
		C3	C2	C3
ClustScan	ACP	KS AT dh KR ACP	KS AT ACP	KS AT dh kr ACP
		C3 R|R	C2	C3 R|x

rifB

	Module	Module	Module
annotated	KS AT DH KR ACP	KS AT dh KR ACP	KS AT dh KR ACP
	C3 x|x	C3 R|S	C3 S|R
SearchPKS	KS AT DH KR ACP	KS AT DH KR ACP	KS AT DH KR ACP
	C3	C3	C3
ClustScan	KS AT DH KR ACP	KS AT dh KR ACP	KS AT DH KR ACP
	C3 R|R	C3 x|S	C3 x|S

rifC

	Module
annotated	KS AT dh KR ACP
	C3 S|S
SearchPKS	KS AT DH KR ACP
	C3
ClustScan	KS AT DH KR ACP
	C3 x|S

rifD

	Module
annotated	KS AT dh KR ACP
	C3 R|R
SearchPKS	KS AT DH KR ACP
	C3
ClustScan	KS AT dh KR ACP
	C3 R|R

rifE

	Module	Module
annotated	KS AT DH KR ACP	KS AT DH KR ACP
	C2 x|x	C3 x|x
SearchPKS	KS AT DH KR ACP	KS AT DH KR ACP
	C2	C3
ClustScan	KS AT DH KR ACP	KS AT DH KR ACP
	C2 R|R	C3 x|S

Legend:

All domain names written in capital letters are active.
Domain names written in small letters are inactive.
Domain names in red differ from annotated clusters.

AT:
C2 - malonyl
C3 - methylmalonyl
CX - not predictable

KR stereochemistry:
Me|OH
R - Rectus
S - Sinister
x - not predictable