NIST Special Database 10 Supplemental Fingerprint Card Data

Craig Watson
Patricia Flanagan

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

NIST Special Database 10 Supplemental Fingerprint Card Data

Craig Watson
Patricia Flanagan
Information Technology Laboratory
Information Access Division

May 2016

U.S. Department of Commerce
Penny Pritzker, Secretary

National Institute of Standards and Technology
Willie May, Under Secretary of Commerce for Standards and Technology and Director

Table of Contents

1.0 Introduction 1
1.1 File Naming Structure 1
1.2 Conversion from ANSI/NIST Data Format 1
1.3 Image File Conversion 1
1.4 Text File Creation 2
1.5 Database Reflectance Calibration 3
2.0 Database File Hierarchy 3
2.1 Low count Loops 4
2.2 NCIC Classifications 4
2.3 Class Referencing 5
2.4 Segmenting the Fingerprint Images 5
2.5 Inked and Live Scan Printed 5
3.0 References 6
4.0 Appendixes 6
Appendix A: Database Fingerprint Image Samples 7
Appendix B: Database Reflectance and Resolution Calibration 9
Appendix C: Fingerprint Class Distribution Statistics 10

NIST Special Database 10
Supplemental Fingerprint Card Data

1.0 Introduction

This document describes the updated web released *NIST Special Database 10*, which contain 552 non-mated fingerprint cards with 5520 fingerprint images with corresponding metadata files per image. The database provides a large sample of patterns for transitional fingerprint classes and classes with a low natural frequency of occurrence in *NIST Special Database 9*. The prints are 832 (w) X 690 (h) at 19.7 pixels per millimeter (500 pixels per inch) (Appendix A). The original CD-ROM released NIST *Special Database 10* was released in a three CD-ROM set in 1993 and contained 5520 fingerprint images in NIST IHEAD raster data format and compressed with JPEG lossless (JPEGL) [1]. The IHEAD header had the metadata that was provided in the ANSI/NIST format (AN2) [2] file and can be found in the metadata txt file included in this new release of the database.

The data was collected by selecting non-mated fingerprint cards, from the FBI's Technical Master File, which contained the most occurrences of the desired fingerprint patterns. Since the entire card was digitized, there is a mix of other classes within the specific class groupings. The data also includes a significant number of referenced fingerprints. Appendix C shows the exact distribution of the classes and referencing for each major class group that was collected.

The specific classes being collected were Tented Arch, Arch, Low Ridge Count Loops, Central Pocket Whorls, Double Loop Whorls, Plain Whorls and Accidental Whorls. The fingerprints are classified using the National Crime Information Center (NCIC) classes [5] assigned by the FBI.

1.1 File Naming Structure

This updated version of SD10 has the same data structure as the original. The directories are named after the original ANSI/NIST an2 file (e.g. card_0001) and will contain ten Portable Network Graphics (PNG)[3] images and corresponding metadata text files.

1.2 Conversion from ANSI/NIST Data Format

The information about each fingerprint image and the fingerprint images were extracted from the AN2 file using NIST Biometric Image Software (NBIS)[4] an2ktool and Netpbm packages. Below are the steps taken to convert the original database images to the updated web version.

1.3 Image File Conversion

1. The NBIS *An2ktool* was used to extract Type-1, Type-2, and Type-14 record information and images from an ANSI/NIST file. Each image file was renamed to reflect the name of the original ANSI/NIST file and the finger position of the image (e.g., card_001_01.png).

2. The NBIS *Djpegl* program was used to convert the JPEGL file into a) a raw image and b) a NISTCOM (NCM) (1) text file which contains the image height and width information.
3. The Netpbm *Rawtopgm* program, along with the height and the width of the image was used to convert the raw image into a PGM (Netpgm grayscale image) file.
4. The Netpbm *Pnmtopng* program was used to convert the PGM to a PNG (e.g., card_001_01.png) file.

1.4 Text File Creation

1. There is no information from the Type-1 record that was extracted for the TXT file.

2. The Type-2 record contains the field (2.024) that has the gender of the subject in the AN2 file. The Gender information was extracted and put into a metadata file with the same naming as the image file but with the TXT extension.

3. The Type-14 record holds the image for the subject and some information pertaining to the image. The field (14.013) has the finger position of the image which is extracted and put in the metadata file. The field (14.020) that has the Class, History and Scan type values e.g., " Class: A, History: aa000009.pct aa tape506.aa001.09 4096x1536, Scan Type: i" is extracted and put into the existing metadata file that corresponds to the image file.

4. Each TXT file contains the Gender, the finger position of the image and the Class, History and Scan type of the image.

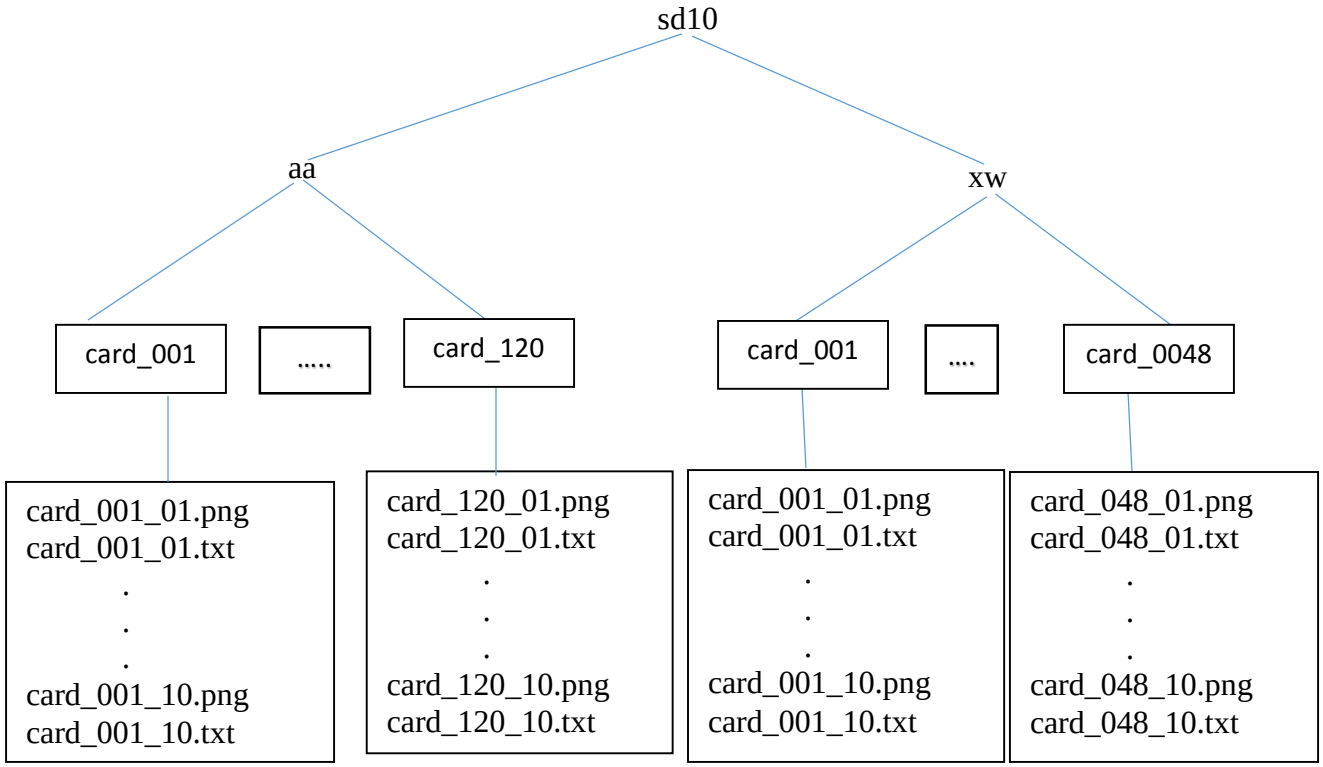

Figure 1: Arrangement of fingerprint files in first and last directories within NIST *Special Database 10*

1.5 Database Reflectance Calibration

The reflectance values for the fingerprint database, NIST *Special Database 10*, were calibrated using a reflection step table [6]. A plot of the reflectance values obtained using this step table is shown in Appendix B. Also shown on the plot and below is an equation used to predict the reflectance of a given data point. The plot in Appendix B shows that this predicted reflectance closely follows the actual reflectance obtained using the reflection step table.

$$\text{predicted \% reflectance} = -5.1 + (.36 * \text{grayscale pixel value})$$

2.0 Database File Hierarchy

The data directory contains the fingerprint images stored in two levels of subdirectories for easier access and clarity (see Figure 1). The first level of subdirectories indicate the specific class collected within that group of data. The specific class groupings: arches (**aa**), central pocket whorls (**cw**), double loop whorls (**dw**), low ridge count loops **(sl**), plain whorls **(pw**), tented arches (**tt**) and accidental whorls (**xw**). The next level has a subdirectory for each fingerprint card which contains the ten segmented fingerprint images for each card along with metadata text files for each image. The **aa**, **sl** and **tt** class groups have 120 cards each (card_001 - card_120) and the other class groups have 48 cards each.

Fingerprints are stored with filenames containing the word card with a "_" then three-digit number followed by a "_" and a two-digit number that represents the finger number and ending in a ".png" extension. There is a text file that accompanies each png file that has the gender of the individual, finger position, class distribution, history and scan type. The two characters in the naming directory are the same as the specific class being collected (**aa**, **tt**, **sl**, **cw**, **dw**, **pw** or **xw**). The first three digits in the image name indicate the sequence number (001 – 120) for **aa**, **sl** and **tt** groups). The finger number is given by the last two digits on the fingerprint card (see Figure 2 for fingerprint card layout). Every ten prints in sequential order are a group of prints from the same card (i.e. digits 1-10).

1. R. Thumb	2. R. Index	3. R. Middle	4. R. Ring	5. R. Little
6. L. Thumb	7. L. Index	8. L. Middle	9. L. Ring	10. L. Little

Figure 2: Layout of fingerprint card numbers.

2.1 Low count Loops

The main low count loops collected were ulnar loops (see statistics in Appendix C). The user can create low count radial loop data by flipping the ulnar loop image about a vertical axis centered on the image. This allows for more samples of both low count loop classes.

2.2 NCIC Classifications

The classes stored in the metadata file (example: Gender: M Position: 09 Class: A History: aa000009.pct **aa** tape506.aa001.09 4096x1536 Scan Type: l) are the NCIC classes, including any references, that were assigned by the FBI Identification Division Automated System (IDAS). A listing of the possible class codes is given below (see Figure 3). Note that the classification ac (approximate class) means that the classification immediately after the ac is the best classification that can be assigned to the print given the information shown in the image. See Figure A.1 (Appendix A) to see the fingerprint image referenced in the above metadata file example.

Classification name (class code)

Arch (**aa**)
Tented Arch (**tt**)
Ulnar Loop Ridge Counts (**01 - 49**)
Radial Loop Ridge Counts (**51 - 99**)
Plain Whorl Inner Ridge Tracing (**pi**)

Plain Whorl Outer Ridge Tracing (**po**)
Plain Whorl Meeting Ridge Tracing (**pm**)
Central Pocket Whorl Inner Ridge Tracing (**ci**)
Central Pocket Whorl Outer Ridge Tracing (**co**)
Central Pocket Whorl Meeting Ridge Tracing (**cm**)
Double Loop Whorl Inner Ridge Tracing (**di**)
Double Loop Whorl Outer Ridge Tracing (**do**)
Double Loop Whorl Meeting Ridge Tracing (**dm**)
Accidental Whorl Inner Ridge Tracing (**xi**)
Accidental Whorl Outer Ridge Tracing (**xo**)
Accidental Whorl Meeting Ridge Tracing (**xm**)
Approximate Class (**ac**) followed by a valid class
Amputation or Missing (**xx**)
Scar or Mutilation (**sr**)

Figure 3: Classification codes for NIST *Special Database 10*.

2.3 Class Referencing

The referencing of a fingerprint is caused by a variety of ambiguities such as a scar occurring in the fingerprint, the quality of the print rolling, or the print having ridge structures characteristic of two different classes. The referenced prints could easily cause a wrong classification when used in testing an automated classification system but could provide a challenge in the later stages of development. The metadata file that accompany the prints in *NIST Special Database 10* contain the primary fingerprint class followed by any references. An example of a fingerprint is shown in Figure A.2 (Appendix A). The fingerprint is classified as an **aa** and referenced to a **tt**.

2.4 Segmenting the Fingerprint Images

The individual fingerprint images were obtained by scanning all ten prints on a card into one large image (4096 X 1536 pixels) and then segmenting each individual image from that larger image. The individual images were segmented at the same exact points on all the larger (full card) images. The image size of 832 X 768 pixels was selected to allow the user the capability of reconstructing the fingerprint card image and re-segmenting the individual fingerprint images if desired. The images overlap by 32 pixels with horizontally adjacent images and by 18 pixels with vertically adjacent images which must be accounted for when reconstructing the fingerprint card image.

2.5 Inked and Live Scan Printed

The fingerprints were scanned from two types of prints. The first type were fingerprint patterns created by rolling the individuals ink covered finger on the fingerprint card (denoted by an '**i**' in the scan type line in the text file. The second type were patterns taken with a live scanning device and then printed onto a fingerprint card (denoted by an '**l**' in the scan type line in the text

file. This is important in that the quality of the image is affected by the resolution of the printer used to print the live scanned image onto the fingerprint card.

3.0 References

1. Mangold Ed., K. *Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information.* s.l. : American National Standard, ANSI/NIST-ITL 1-2011:2015.

2. *WSQ Gray-scale Fingerprint Image Compression Specification, IAFIS-IC-0110v2,*. Federal Bureau of Investigation,. s.l. : Federal Bureau of Investigation,, February 1993.

3. K. Mangold Ed. *Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information.* s.l. : American National Standard, ANSI/NIST-ITl 1-2011:2015.

4. *Information Technology -- Computer graphics and image processing -- Portable Network Graphics (PNG) Functional specificaiton.* s.l. : International Standards Organization/International Electrotechnical Commission, 2004.

5. *"Information Technology -- Lossless and near-lossless compression of continuous tone still images: Baseline".* s.l. : International Standards Organizaiton/International Electrotechnical Commission, ISO/IEC 14495-1, 1999.

6. C. Watson et al. *"User's Guide to NIST Biometric Image Software,".* s.l. : National Institute of Standards and Technology ITL/IAD/Image Group, 2004.

7. *The Science of Fingerprints.* Washington, DC : U.S. Department of Justice, 1984.

8. *Standard Reference Materials Reflection Step Table 2601*. s.l. : National Bureau of Standards.

4.0 Appendixes

The Appendixes are from the original *NIST Special Database 10* documentation. The original database that became public was distributed in 1993 as a three disk set. The information in the Appendixes reflect the original data.

Appendix A: Database Fingerprint Image Samples
Appendix B: Database Reflectance and Resolution Calibration
Appendix C: Fingerprint Class Distribution Statistics

Appendix A: Database Fingerprint Image Samples

Figure A.1: Fingerprint file **aa000001.pct** from *NIST Special Database 10.*

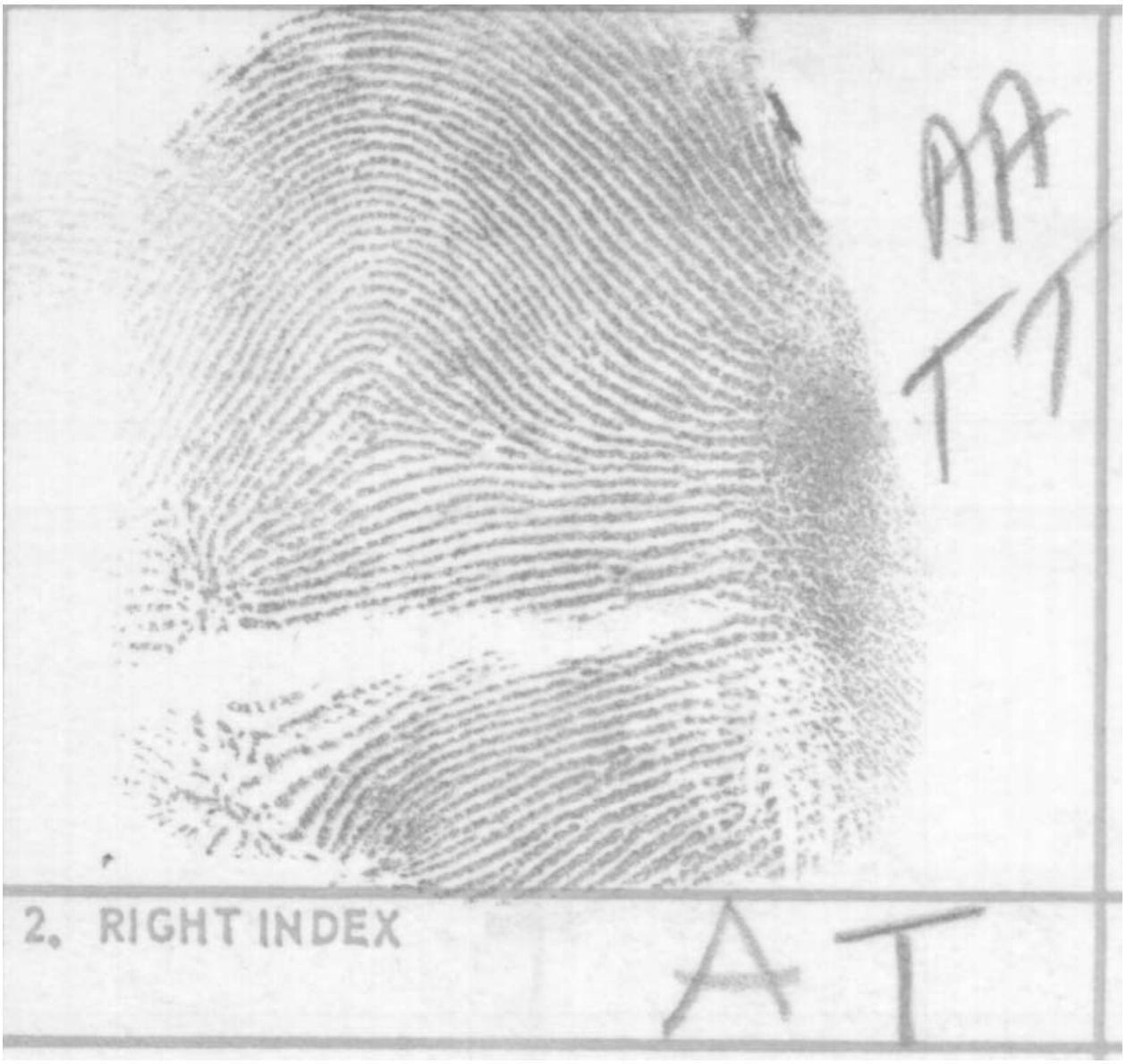

Figure A.2: Fingerprint file **aa000042.pct** from NIST *Special Database 10.*

Appendix B: Database Reflectance and Resolution Calibration

Contact Craig Watson
for Reflectance Plot
e-mail => craig.watson@nist.gov
phone => (301) 975-4402

Database Resolution Calibration

The resolution of the scanner used to create the database was calibrated using a NBS 1010A resolution chart[1]. A portion of this image is shown below (Figure B.1) after being magnified for viewing (Note: Printing has significantly reduced the quality of this image.). The scan of the table showed that the vertical resolution was approximately 10.5 line pairs per millimeter and the horizontal resolution was approximately 11.0 line pairs per millimeter.

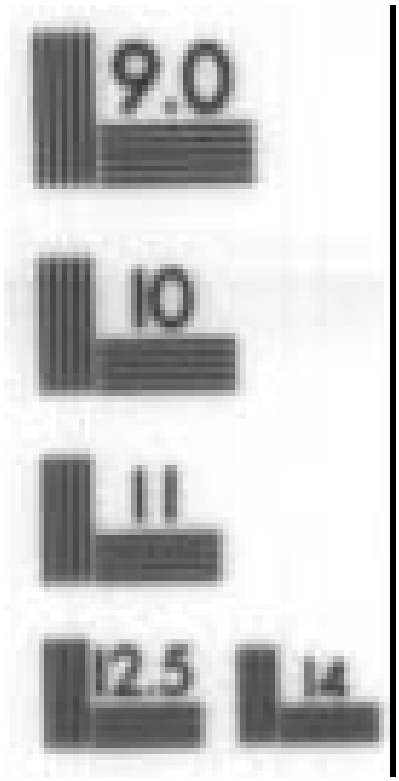

Figure B.1: Scan of Resolution Calibration Table.

1. National Bureau of Standards, Microcopy Resolution Test Chart, Standard Reference Material 1010a, ANSI and ISO test chart No. 2.

Appendix C: Fingerprint Class Distribution Statistics

Class Distribution

The data in Figure C.1 shows the exact class distribution of *NIST Special Database 10*. Since the data was collected by storing all ten fingerprints from a card some of the specific class groupings contain numerous fingerprints from other classes.

Note:

?L (1-5 ridge count loops)
?M (6-30 ridge count loops)
?H (>30 ridge count loops)

All whorl ridge tracings (I,M,O) for a whorl class are counted together (i.e. PI,PM,PO are counted in PW).

Specific Class Groupings

Actual Classes Collected	aa	sl	tt	cw	dw	pw	xw	total
AA	1199		88					1287
UL		843	149			2		994
UM		164	86			71		321
RL		189	23					212
RM			10					10
TT			842					842
PW				153	138	228	250	769
CW				308		94	7	409
DW				19	342	79	124	564
XW						6	99	105
SR	1	4	2					7
Total	1200	1200	1200	480	480	480	480	5520

Figure C.1: Class distribution statistics for NIST *Special Database 10*.

Class Referencing

The data in Figure C.2 shows the class referencing for *NIST Special Database 10*. The referencing statistics are given for each specific class grouping. All primary classes without references (UL, UM, ...) show the number of unreferenced prints for that particular class. See the note on page 21 for an explanation of the L, M, and H, as used with the ulnar and radial loops. If a primary class has more than one reference the ordering is insignificant, meaning UL/AA/TT is the same as UL/TT/ AA and both would be counted as UL/AA/TT.

Figure C.2: Class referencing statistics for NIST *Special Database 10*

Specific Class Groupings

Class/Reference	aa	sl	tt	cw	dw	pw	xw	total
UL		812	178			69		1059
UL/UM		5	5					10
UL/TT		184	48					232
UL/UM/TT		1	3					4
UL/PW/DW			1					1
UL/CW/TT		1						1
UL/AA/TT		2						2
UM UM/UL		3	3					6
UM/CW						3		3
UM/TT			3					3
UM/UL/TT			1					1
UM/CW/DW						1		1
RL		115	18					133
RL/TT		70	9					79
RL/RM/TT			1					1
RM RM/RH			1					1
RM/CW			1					1
RM/TT			2					2
RM/PW/CW			1					1
RM/DW/XW			1					1
PW				145	120	198	220	683
PW/CW				8		9	4	21
PW/DW					18	21	26	65

Specific Class Groupings

Class/Reference	aa	sl	tt	cw	dw	pw	xw	total
CW				216		81	1	298
CW/UM				10		5	2	17
CW/RM				6			1	7
CW/PW				70		8	1	79
CW/DW				2			1	3
CW/PW/DW				1			1	2
CW/PW/XW				1				1
CW/DW/XW				2				2
DW				11	268	67	105	451
DW/UM				1	3	1		5
DW/PW				1	65	10	16	92
DW/CW				5	4	1	1	11
DW/XW							1	1
DW/PW/CW					2		1	3
XW						5	77	82
XW/RM							1	1
XW/PW							3	3
XW/CW							2	2
XW/DW						1	13	14
XW/PW/DW							3	3
AA	1160		68					1228
AA/TT	32		20					52
AA/UL/TT	1							1
TT			341					341
TT/UL			310					310
TT/UM			13					13
TT/RL			55					55
TT/RM			4					4
TT/CW			1					1
TT/AA			77					77
TT/UL/UM			6					6
TT/UL/RL			3					3
TT/UL/AA			22					22
TT/RL/RM			1					1
TT/RL/AA			2					2
Total scars	7	15	10	1				33