

# Proposal to encode MALAYALAM SIGN CIRCULAR VIRAMA

Cibu Johny, [cibu@google.com](mailto:cibu@google.com)  
 Shiju Alex, [shijualexonline@gmail.com](mailto:shijualexonline@gmail.com)  
 Sunil V S, [vssun9@gmail.com](mailto:vssun9@gmail.com)  
 2014-Jan-09

## Introduction

The Virama in Malayalam script is called *Candrakkala* in its explicitly visible form: ഓ. It has got the following two main functions:

1. Represent the neutral vowel sound (ə) at the word endings. This sound is called *Samvrtōkārām* and transliterated as ũ.<sup>1</sup>
2. Remove the inherent vowel A from a Malayalam consonant letter to indicate the corresponding pure consonant.

*Candrakkala* was introduced in Malayalam script, by Rev. Dr. Hermann Gundert around 1847 to represent *Samvrtōkārām*. Before that, *Samvrtōkārām* was implicit at word endings with nominal consonant letters and often contextual as well. After around half a century, by around 1900, it acquired the second function to represent pure consonants without their inherent vowel A.

During this half century, before *Candrakkala* being used for its second function, Malayalam script tried a separate symbol, that is a small circle at the top-right corner of the consonant letter, to represent the pure consonant.



This symbol co-existed with *Candrakkala* while *Candrakkala* doing its original function of neutral vowel sound. At the turn of the 20th century, circular virama got deprecated and *Candrakkala* was overloaded with the second function as well. Circular virama must be encoded to record the documents from its usage period.

## Relation to the Combining Anusvara Above proposal

The document L2/14-003<sup>2</sup> proposes a similar looking circle sign for Combining Anusvara Above. The Circular Virama and Anusvara Above differ in their placement with respect to the consonant letter. Anusvara Above is placed at the top-center position of the letter, while

<sup>1</sup> As per ISO 15919 described in <http://homepage.ntlworld.com/stone-catend/trind.htm>

<sup>2</sup> <http://www.unicode.org/L2/L2014/14003-malayalam-anusvara-above.pdf>

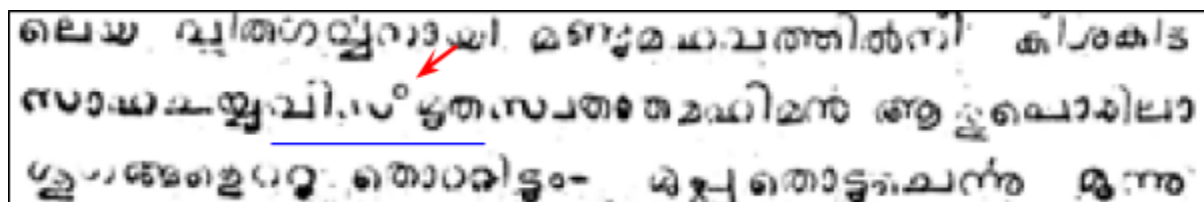
The attestations of Anusvara Above and Circular Virama are from two different era. Anusvara Above attestations are around 400 years old while Circular virama is at the most 150 years old. These characters are used for unrelated functions as well.

This letter could be confused with the degree sign (00B0). At small font sizes, it could be confused with Virama(0D4D). To avoid such security issues, this character should be excluded from IDNs since it is not needed for contemporary Malayalam use.

Attestations from various sources are listed below as sources organized as subsections.

foreign words which ended in certain consonants, as വാക്ക, 'വാക്,' മനസ്സ, 'മനസ്,' സമ്പത്ത, 'സമ്പത്,' ഭാവീക

### Khāṇḍvadāham (1867), by Kaṭattanāṭṭu Śaṅkaravarmma<sup>4</sup>



<sup>3</sup> <https://drive.google.com/file/d/0B4hBi4hJ-5mnTTdTbWJ0VUJyWE0/edit?usp=sharing> page 11

<sup>4</sup> [https://archive.org/stream/KhandavaDahamKadhakali/Khandava\\_Daham\\_Kadhakali#page/n20/mode/1up](https://archive.org/stream/KhandavaDahamKadhakali/Khandava_Daham_Kadhakali#page/n20/mode/1up) page 20

## First edition of Manorama Newspaper (1890)<sup>5</sup>

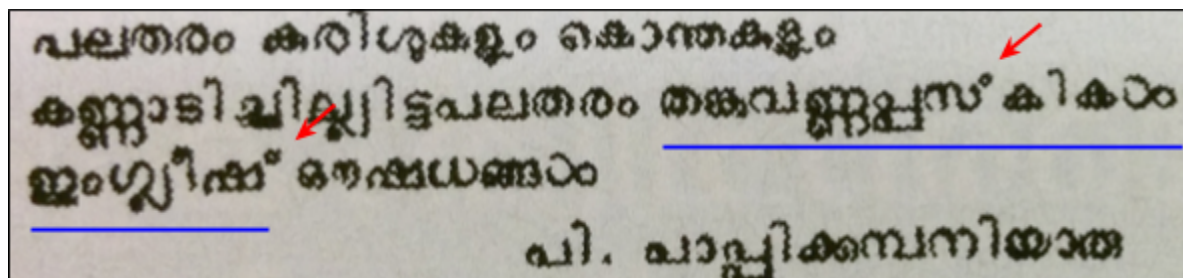


Image 3: The underlined words are തങ്കവണ്ണപ്പസ്കികൾ (taṅkavaṇṇappaskikal), ഇംഗ്ലീഷ് (iṅḡlīṣ). This document uses vowel sign U to represent *Ṣaṁvṛtōkārāṇ*.

## Kavibhāratam (1893) by Koṭṭūṇhallūr Kuñṅikkuṭṭan Taṁpurān<sup>6</sup>

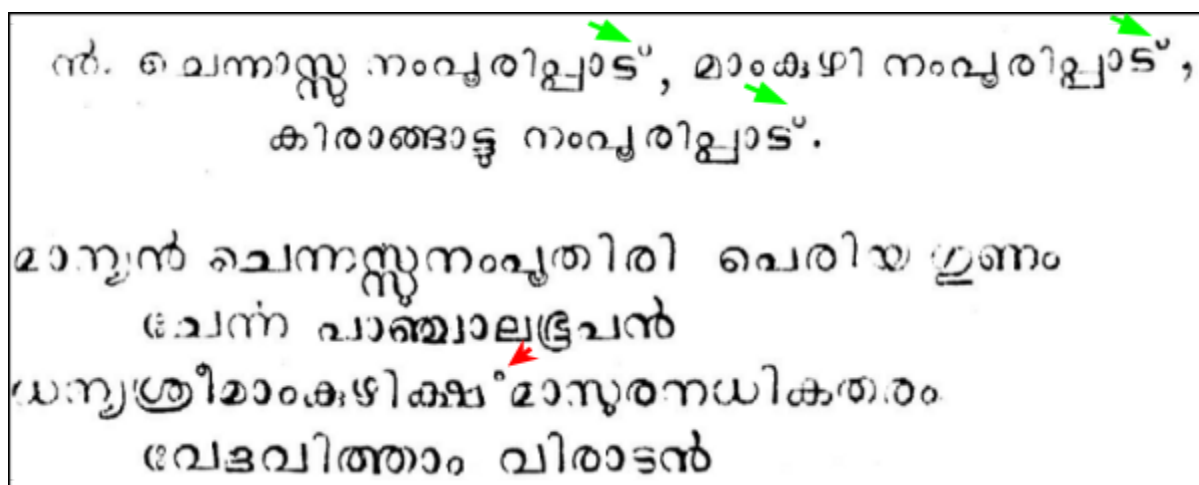


Image 4: Contrasting usage of *Candrakkala* and circular virama. First 3 usages are of *Candrakkala* for *Ṣaṁvṛtōkārāṇ*. The last one is of Circular Virama for vowelless consonant to split up the grapheme cluster K-SS-MA.

<sup>5</sup> <https://lh6.googleusercontent.com/-FgZuIN0h9G4/Urilyx2Jcbl/AAAAAAAAAEvi4/nSWYhKbfJC0/w1912-h0-no/13+-+1>  
page 20

<sup>6</sup> [http://www.keralasahityaakademi.org/online\\_library/m-k.html](http://www.keralasahityaakademi.org/online_library/m-k.html) page 5, 6

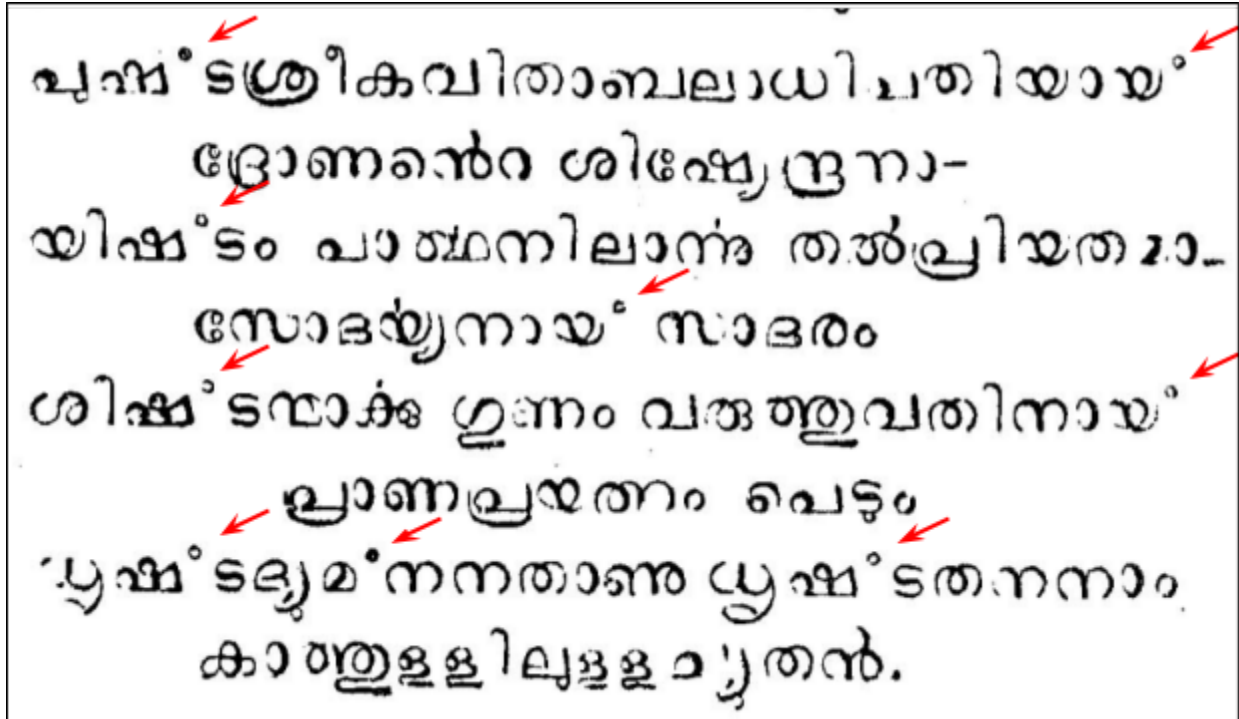


Image 5: All usages are of Circular Virama for pure consonant.

**Aṃbarīṣacaritaṃ(1897) by Kaṭattanāṭṭu Udayavarmma<sup>7</sup>**

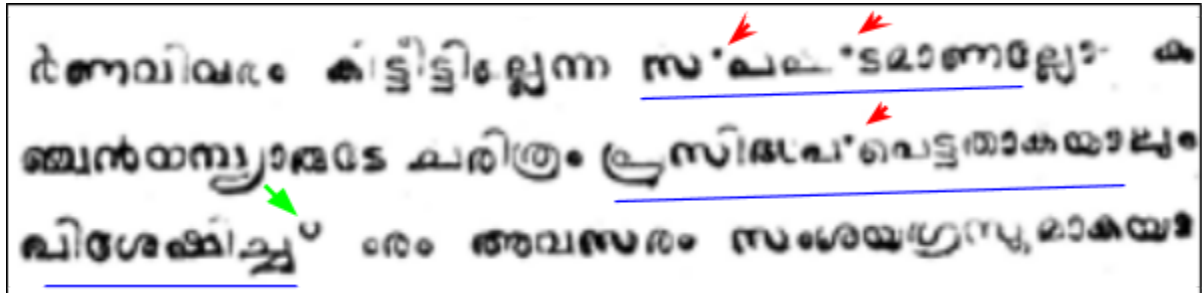


Image 6: Contrasting usage of *Candrakkala* and circular virama. The words that use circular virama for pure consonant:

- സ<circular virama>പഷ<circular virama>മാണല്ലോ (*spaṣṭamāṇallō*)
  - പ്രസിദ്ധപ<circular virama>പെട്ടതാകയാലും(*prasiddhappeṭṭatākayāluṃ*) for
- The last word that use *Candrakkala* for Saṃvṛtōkāraṃ is വിശേഷിച്ച് (*viśēṣiccū*)

<sup>7</sup> [https://archive.org/stream/AmbareeshaCharithamThullal/Ambareesha\\_Charitham\\_Thullal#page/n3/mode/1up](https://archive.org/stream/AmbareeshaCharithamThullal/Ambareesha_Charitham_Thullal#page/n3/mode/1up)  
pages 1 and 9

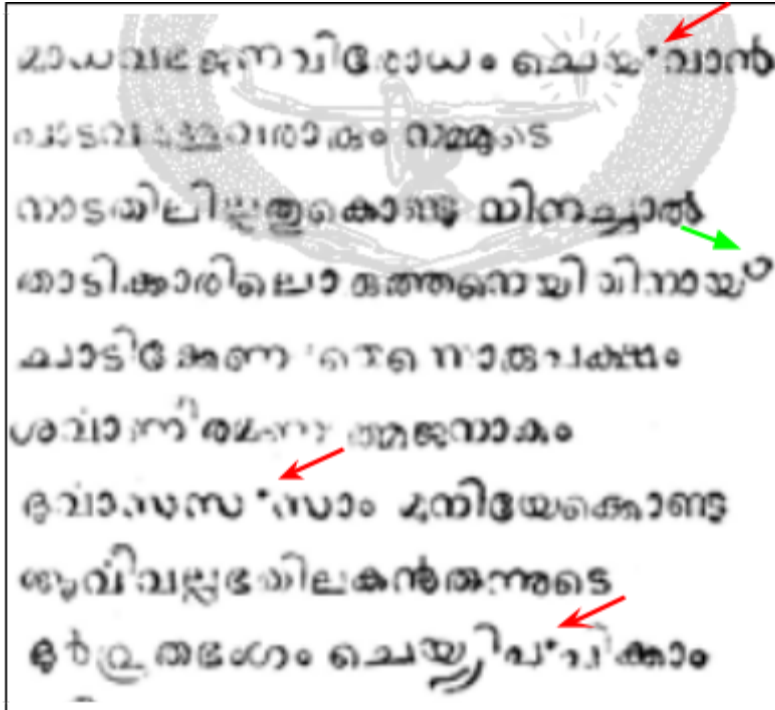


Image 7: Another instance of contrasting usage of *Candrakkala* and circular virama. Second one is *Candrakkala* for *Samvrtōkārṇ* at the word ending. The first, third and fourth are circular viramas, used to indicate vowelless consonants, so as to split up the grapheme clusters. *Candrakkala* is bigger and circular virama is tiny.

### Indyayile pāṣika|(1903) by Kaṇṭattil Māttu||a Māppiḷa<sup>8</sup>

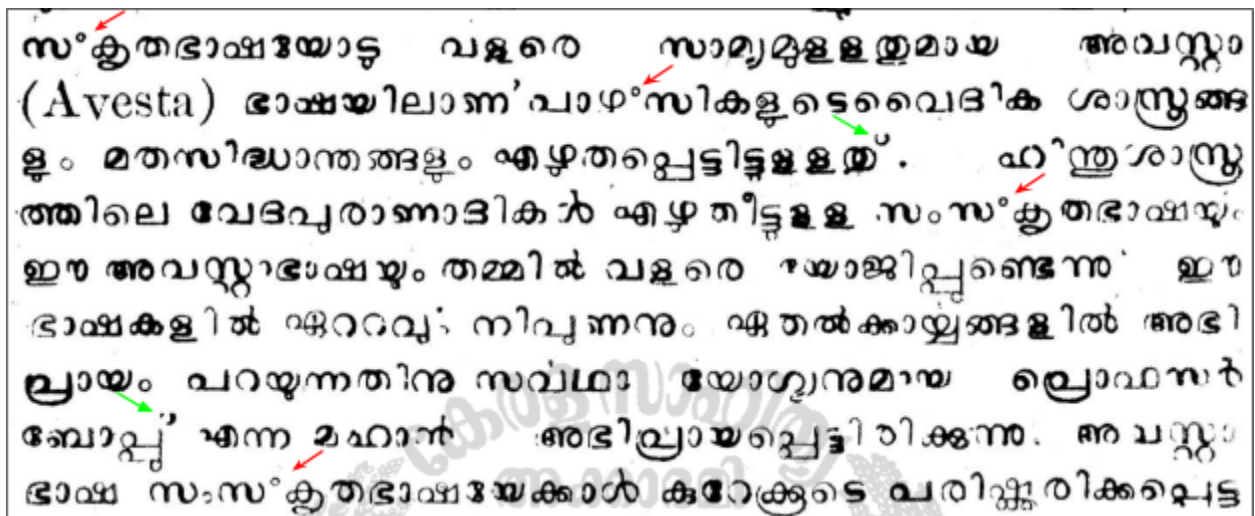


Image 8: An instance of contrasting usage of *Candrakkala* and Circular Virama. Third and the fifth ones marked by green arrows from the left are *Candrakkala* for *Samvrtōkārṇ* at the word endings. The first, second, fourth and sixth are Circular Virama for vowelless consonants to split up the grapheme clusters.

<sup>8</sup> [http://www.keralasahityaakademi.org/online\\_library/m-i.html](http://www.keralasahityaakademi.org/online_library/m-i.html) page 9

### Kaṇṭhāmṛtam (1906) by Nīlakaṇṭha Tīrtthar

കാരമെന്ന അർത്ഥത്തിലാണമയ (ഃ) പ്രയോഗമെങ്കിൽ. ൫. അ  
 ൧൩ ആനന്ദത്തിനു കാരണമെന്നു പറയുന്നതിനും. ൬. സംസ്  
 യ ജീവനല്ല. ൭. ജീവനും ആനന്ദമയനും.  
 ൧. സോകാമയതയെന്ന മന്ത്രത്തിൽ ആനന്ദമയൻ കായ്യി  
 ന്നു വരുന്നതിനാലെന്നുസാരം. ൨. അനുമാനം. ൩. തന്റെ  
 പരമനായജീവനും. ൪. പാപാഭിയിക്തിനിന്നും മേൽനില്ക്കയാൽ  
 ഉദിതിനാമ' എന്ന വാക്യത്തിൽ 'ഉദേ' എന്ന പേർ മുതലായു

Image 9

Based on the above attestations we can see that this character was in use in the later part of the 19<sup>th</sup> century and the early years of the 20<sup>th</sup> century. Since then it has been replaced by the Virama: ॐ (0D4D).

## Choice of codepoint

The proposed codepoint is **0D3C** which is a location close to Virama (0D4D).

## Malayalam Code Chart

	0D0	0D1	0D2	0D3	0D4	0D5	0D6	0D7
0	◌̇	ഐ	ഓ	ര	ീ		ഋ	ൠ
1	◌̈		ഡ	റ	ു		ൡ	ൢ
2	◌̉	ഒ	ഘ	ല	ൂ		ൣ	൤
3	◌̊	ഓ	ണ	ള	്യ		൥	൦
4		ഔ	ത	ഴ	ൂ	൦		൧
5	അ	ക	ഥ	വ		യ	ൡ	ൢ
6	ആ	ഖ	ഭ	ശ	െ	ഴ	ഓ	ഘ
7	ഇ	ഗ	ധ	ഷ	േ	ൠ	ൡ	ഘ
8	ഈ	ഘ	ന	സ	ൈ	ധ	ൣ	൤
9	ഉ	ങ	ണ	ഹ		൦	ൢ	൤
A	ഊ	ച	പ	ട	ൊ	ൡ	ൣ	൥
B	ഋ	ൠ	ഫ	◌̇	ോ	ൢ	ൣ	൤
C	ൡ	ജ	ബ	◌̈	ൠ	ൢ	ൣ	൤
D		ഡ	ഭ	ഌ	ൠ	ൢ	ൣ	൤
E	എ	ണ	മ	ാ	ൠ	ൢ	ൣ	൤
F	ഏ	ട	യ	ി	ൠ	ൢ	ൣ	൤

## Legend

*White:* Published as of Unicode 6.3.

*Green:* Accepted and at advanced stage not permitting changes.

*Yellow:* Fractions, accepted.

*Orange:* Chillu LLL accepted but preferably to be moved to group with Chillu Y.

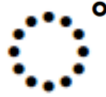
*Blue:* New location proposed for Chillu LLL.

*Purple:* Locations of ongoing proposals of Anusvara Above, Chillus M and Y, Vertical Bar Virama and Symbol Para.

*Red:* Location proposed for Circular virama.



## Proposed character with properties



OD3C;MALAYALAM SIGN CIRCULAR VIRAMA;Mc;0;L;;;;;N;;;;;

## Official Proposal Summary Form

(Based on N3902-F)

### A. Administrative

1. Title

**Proposal to encode MALAYALAM SIGN CIRCULAR VIRAMA**

2. Requester's name

**Cibu Johny, Shiju Alex, Sunil V S**

3. Requester type (Member body/Liaison/Individual contribution)

**Individual contribution**

4. Submission date

**2014-Jan-09**

5. Requester's reference (if applicable)

6. Choose one of the following: This is a complete proposal (or) More information will be provided later

**This is a complete proposal.**

### B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters), Proposed name of script

**No**

1b. The proposal is for addition of character(s) to an existing block, Name of the existing block

**Yes, Malayalam**

2. Number of characters in proposal

**1 (one)**

3. Proposed category

**Category B1, specialized small (for this character)**

4. Is a repertoire including character names provided?

**Yes**

4a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

**Yes**

4b. Are the character shapes attached in a legible form suitable for review?

**Yes**

5. Fonts related:

a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

**Michael Everson**

b. Identify the party granting a license for use of the font by the editors (include address, e-mail etc.)

**Michael Everson, everson@evertyp.com**

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

**Yes**

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?



**Yes**

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

**Yes**

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

**See detailed proposal.**

### C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

**No**

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

**Yes**

2b. If YES, with whom?

**The authors themselves are part of the user community.**

2c. If YES, available relevant documents

**None specifically. The matter was discussed via email.**

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

**Those (scholars and enthusiasts) who desire to store old Malayalam documents using this character.**

4a. The context of use for the proposed characters (type of use; common or rare)

**Rare**

4b. Reference

**See detailed proposal.**

5a. Are the proposed characters in current use by the user community?

**No. This character was in historical use.**

5b. If YES, where?

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

**Yes**

6b. If YES, is a rationale provided?

**It belongs in the Malayalam block which is in the BMP.**

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

**Only one character is proposed.**

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

**No**

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

**No**

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

**No**

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences?

**No**

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

**No.**

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

**No**

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

13c. If YES, reference: